



TUGAS AKHIR - SS141501

**PEMODELAN FAKTOR-FAKTOR
YANG MEMENGARUHI JUMLAH KASUS TUBERKULOSIS
DI JAWA TIMUR MENGGUNAKAN REGRESI
NONPARAMETRIK *SPLINE***

**FIRDA FAHRUN NISA'
NRP 1312 100 027**

**Dosen Pembimbing
Prof. Dr. I Nyoman Budiantara, M.Si**

**PROGRAM STUDI S1
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2016**



FINAL PROJECT - SS141501

**MODELING FACTORS THAT AFFECT THE TOTAL
NUMBER OF TUBERCULOSIS IN EAST JAVA
USING NONPARAMETRIC *SPLINE* REGRESSION**

**FIRDA FAHRUN NISA'
NRP 1312 100 027**

**Supervisor
Prof. Dr. I Nyoman Budiantara, M.Si**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2016**

LEMBAR PENGESAHAN
PEMODELAN FAKTOR-FAKTOR
YANG MEMENGARUHI JUMLAH KASUS
TUBERKULOSIS DI JAWA TIMUR
MENGGUNAKAN REGRESI NONPARAMETRIK SPLINE

TUGAS AKHIR

Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada
Program Studi S-1 Jurusan Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember

Oleh :
FIRDA FAHRUN NISA'
NRP. 1312 100 027

Disetujui oleh Pembimbing Tugas Akhir :

Prof. Dr. I Nyoman Budiantara, M.Si
NIP. 19650603 198903 1 003

[Handwritten signature]
(03)

Mengetahui
Ketua Jurusan Statistika FMIPA-ITS



Dr. Suhartono
NIP. 19710929 199512 1 001

SURABAYA, JULI 2016

**PEMODELAN FAKTOR-FAKTOR
YANG MEMPENGARUHI JUMLAH KASUS
TUBERKULOSIS DI JAWA TIMUR MENGGUNAKAN
REGRESI NONPARAMETRIK *SPLINE***

Nama Mahasiswa : Firda Fahrún Nisa'
NRP : 1312 100 027
Jurusan : Statistika FMIPA-ITS
Dosen Pembimbing : Prof. Dr. I Nyoman Budiantara, M.Si

ABSTRAK

Tuberkulosis adalah penyakit menular langsung yang disebabkan oleh kuman TB (Mycrobacterium tuberculosis), sebagian besar kuman TB menyerang paru, tetapi dapat juga mengenai organ tubuh lainnya. Secara nasional 7,5% angka kematian di Indonesia disebabkan oleh penyakit tuberkulosis. Provinsi Jawa Timur memiliki kasus TB terbanyak kedua setelah provinsi Jawa Barat. Data Dinas Kesehatan Provinsi Jawa Timur tahun 2014 menunjukkan kasus TB mencapai 42.458 kasus. Tindakan prefentif yang dilakukan oleh pemerintah adalah dengan mengendalikan faktor-faktor yang mempengaruhi penyakit TB. Pada penelitian ini, metode regresi nonparametrik spline digunakan untuk memodelkan jumlah kasus TB di Jawa Timur. Metode regresi nonparametrik Spline digunakan karena pola data antara jumlah kasus TB di Jawa Timur tidak menunjukkan kecenderungan pola tertentu dengan faktor-faktor yang diduga mempengaruhinya, yaitu persentase keluarga dengan kepemilikan sarana sanitasi dasar sehat, persentase penduduk miskin, persentase gizi buruk, persentase tenaga kesehatan terlatih dan jumlah rumah tangga ber-PHBS. Berdasarkan hasil pemodelan didapatkan model spline terbaik yaitu spline dengan kombinasi knot. Nilai koefisien determinasi yang didapatkan adalah 83,42% dengan tiga variabel prediktor dalam model berpengaruh signifikan.

Kata Kunci : Jumlah kasus TB, GCV, Titik Knot, Nonparametrik, Spline.

MODELING FACTORS THAT AFFECT THE TOTAL NUMBER OF TUBERCULOSIS IN EAST JAVA USING NONPARAMETRIC *SPLINE* REGRESSION

Name : Firda Fahrur Nisa'
NRP : 1312 100 027
Department : Statistics FMIPA-ITS
Supervisor : Prof. Dr. I Nyoman Budiantara, M.Si

ABSTRACT

Tuberculosis is an infectious disease directly caused by bacteria (Mycobacterium tuberculosis), most of the bacteria attack lungs, but can also on other organs. Nationally 7,5 perrcent mortality in Indonesia are caused by tuberculosis. East Java has the second highest number of tuberculosis cases right after West Java. Data from Health Office of East Java in 2014 show that the number of TB cases reached 42.458 cases. Preventive action by the goverment is by control the factors that influence tuberculosis. In this research, nonparametric spline regression method is used for modeling the number of TB cases in East Java. This method is used because the data patterns between the number of TB cases in East Java showed no inclination specific pattern with the factors that suspected to influence, that is the percentage of families with primary healthy sanitation ownership, the percentage of the poor population, the percentage of malnutrition people, the percentage of trained health workers, and the percentage of households with health and clean behaviors. Based on the generated model, the best spline models is spline with combination knots. Determination coefficient value is 83,42 percent with three predictor variables in the model have a significant effect.

Key words : Number of Tuberculosis cases, GCV, Knots Point, Nonparametric Spline.

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
COVER PAGE	ii
LEMBAR PENGESAHAN	iii
ABSTRAK	v
ABSTRACT	vii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xv
DAFTAR LAMPIRAN	xvii
 BAB I PENDAHULUAN	
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah	4
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	5
1.5 Batasan Penelitian.....	5
 BAB II TINJAUAN PUSTAKA	
2.1 Analisis Regresi.....	7
2.2 Regresi Nonparametrik Spline	8
2.3 Pemilihan Titik Knot Optimal	9
2.4 Pengujian Estimasi Parameter	9
2.5 Kriteria Kebaikan Model	11
2.6 Pemeriksaan Asumsi Residual Dalam Model Regresi..	11
2.7 Tuberkulosis	13
2.8 Penelitian Sebelumnya.....	14
 BAB III METODOLOGI PENELITIAN	
3.1 Sumber Data	17
3.2 Variabel Penelitian.....	17
3.3 Definisi Operasional Variabel	18
3.4 Langkah Analisis dan Diagram Alir	20

BAB IV ANALISIS DAN PEMBAHASAN

4.1	Karakteristik Jumlah Kasus Tuberkulosis di Jawa Timur	23
4.2	Analisis Faktor-Faktor yang Diduga Mempengaruhi Jumlah Kasus TB di Jawa Timur	27
4.3	Pemilihan Titik Knot Optimum	30
4.3.1	Pemilihan Titik Knot dengan Satu Titik Knot	30
4.3.2	Pemilihan Titik Knot dengan Dua Titik Knot	31
4.3.3	Pemilihan Titik Knot dengan Tiga Titik Knot	33
4.3.4	Pemilihan Titik Knot dengan Kombinasi Knot	34
4.3.5	Pemilihan Model Terbaik	35
4.3.6	Penaksiran Parameter Model Regresi Nonparametrik Spline	36
4.4	Pengujian Signifikansi Parameter Model Regresi Nonparametrik Spline	36
4.4.1	Uji Serentak	37
4.4.2	Uji Individu	37
4.5	Pengujian Asumsi Residual	39
4.5.1	Pemeriksaan Asumsi Identik	39
4.5.2	Pemeriksaan Asumsi Independen	40
4.5.3	Pemeriksaan Asumsi Distribusi Normal	41
4.6	Nilai Koefisien Determinasi (R^2)	41
4.7	Interpretasi Model Regresi Nonparametrik Spline	42

BAB V KESIMPULAN DAN SARAN

5.1	Kesimpulan	47
5.2	Saran	48

DAFTAR PUSTAKA

LAMPIRAN

DAFTAR TABEL

	Halaman
Tabel 2.1 <i>Analysis of Variance</i>	10
Tabel 3.1 Variabel Penelitian.....	18
Tabel 3.2 Struktur Data Variabel Penelitian	18
Tabel 4.1 Karakteristik Jumlah Kasus Tuberkulosis dan Faktor-Faktor yang Mempengaruhi	23
Tabel 4.2 Nilai GCV Dengan Satu Titik Knot.....	31
Tabel 4.3 Nilai GCV Dengan Dua Titik Knot	32
Tabel 4.4 Nilai GCV Dengan Tiga Titik Knot.....	33
Tabel 4.5 Nilai GCV Dengan Kombinasi Titik Knot	35
Tabel 4.6 Nilai GCV Minimum Setiap Pemilihan Titik Knot Optimal	36
Tabel 4.7 Analisis Ragam Uji Serentak.....	37
Tabel 4.8 Hasil Uji Individu	38
Tabel 4.9 Analisis Ragam Uji Glejser	40

DAFTAR GAMBAR

	Halaman
Gambar 3.1 Diagram Alir Penelitian	21
Gambar 4.1 Jumlah Kasus Tuberkulosis di Jawa Timur	24
Gambar 4.2 <i>Scatterplot</i> Persentase Keluarga dengan Kepemilikan Sarana Sanitasi Dasar Sehat	27
Gambar 4.3 <i>Scatterplot</i> Persentase Penduduk Miskin	28
Gambar 4.4 <i>Scatterplot</i> Persentase Gizi Buruk Masyarakat	29
Gambar 4.5 <i>Scatterplot</i> Persentase Tenaga Kesehatan Terlatih TB	29
Gambar 4.6 <i>Scatterplot</i> Persentase Rumah Tangga Ber- PHBS	30
Gambar 4.7 <i>Scatterplot</i> antara Residual dengan \hat{y}	39
Gambar 4.8 Plot ACF Residual	40
Gambar 4.9 Hasil Uji <i>Kolmogorov-Smirnov</i>	41
Gambar 4.10 Peta Persebaran Jumlah Kasus TB Berdasarkan Persentase Keluarga dengan Kepemilikan Sarana Sanitasi Dasar Sehat	42
Gambar 4.11 Peta Persebaran Jumlah Kasus TB Berdasarkan Persentase Penduduk Miskin	43
Gambar 4.12 Peta Persebaran Jumlah Kasus TB Berdasarkan Persentase Gizi Buruk Masyarakat	44
Gambar 4.13 Peta Persebaran Jumlah Kasus TB Berdasarkan Persentase Tenaga Kesehatan Terlatih TB	45
Gambar 4.14 Peta Persebaran Jumlah Kasus TB Berdasarkan Persentase Rumah Tangga Ber- PHBS	46

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Data Jumlah Kasus Tuberkulosis di Jawa Timur dengan Faktor-Faktor Yang Mempengaruhi Tahun 2014	51
Lampiran 2 Program Pemilihan Titik Knot Optimal dengan Satu Titik Knot Menggunakan <i>Software R</i>	53
Lampiran 3 Program Pemilihan Titik Knot Optimal dengan Dua Titik Knot Menggunakan <i>Software R</i>	56
Lampiran 4 Program Pemilihan Titik Knot Optimal dengan Tiga Titik Knot Menggunakan <i>Software R</i>	59
Lampiran 5 Program Pemilihan Titik Knot Optimal dengan Kombinasi Titik Knot Menggunakan <i>Software R</i>	62
Lampiran 6 Program Kombinasi Titik Knot 3^5	69
Lampiran 7 Program Estimasi Parameter dengan Kombinasi Titik Knot 1,3,2,3,2.....	70
Lampiran 8 Program Uji Glejser untuk Kombinasi Titik Knot 1,3,2,3,2.....	73
Lampiran 9 Output Nilai GCV dengan Satu Titik Knot.....	75
Lampiran 10 Output Nilai GCV dengan Dua Titik Knot	76
Lampiran 11 Output Nilai GCV dengan Tiga Titik Knot.....	77
Lampiran 12 Output Estimasi Parameter dan Uji Signifikansi Parameter Model.....	78
Lampiran 13 Output Uji Glejser	81
Lampiran 14 Output Uji <i>Kolmogorov-Smirnov</i>	81
Lampiran 15 Surat Pernyataan Data Sekunder	82

BAB I

PENDAHULUAN

1.1 Latar Belakang

Tuberkulosis (TB) di Indonesia merupakan salah satu masalah kesehatan masyarakat yang utama, tergolong penyakit menular dan langsung mengenai parenkim paru yang disebabkan oleh basil *Mycrobacterium tuberculosis*. Kuman TB bersifat tahan asam pada pewarnaan, sehingga dinamakan BTA (Basil Tahan Asam). Sifat kuman TB cepat mati dengan sinar matahari langsung tetapi masih dapat bertahan hidup beberapa jam di tempat yang gelap dan lembap. Sebagian besar kuman TB menyerang paru, tetapi dapat juga menyerang organ tubuh lainnya (Departemen Kesehatan Republik Indonesia, 2001). TB menyebar melalui batuk dan bersin. Proses penularan terjadi ketika seorang yang memiliki TB aktif batuk atau bersin hingga menyebarkan kuman ke udara (Departemen Kesehatan Republik Indonesia, 2009).

Laporan TB dunia oleh WHO tahun 2006, menempatkan Indonesia sebagai penyumbang terbesar nomor tiga di dunia setelah India dan Cina dengan jumlah kasus baru sekitar 539.000 jiwa dengan jumlah 101.000 jiwa per tahun. Di Indonesia penyakit TB merupakan penyakit infeksi yang masih menjadi masalah utama kesehatan masyarakat. Secara nasional 7,5% angka kematian di Indonesia disebabkan oleh penyakit TB (Departemen Kesehatan Republik Indonesia, 2009). Menurut *Global Tuberculosis Control* tahun 2011 jumlah pasien tuberkulosis di Indonesia merupakan ke-4 terbanyak di dunia setelah India, Cina dan Afrika Selatan (Siagian, 2011).

Provinsi Jawa Timur memiliki kasus TB terbanyak kedua setelah provinsi Jawa Barat (Kementrian Kesehatan RI, 2011). Data Dinas Kesehatan Provinsi Jawa Timur tahun 2011 menunjukkan kasus TB mencapai 41.404 kasus, sementara Jawa Barat mencapai 62.563 kasus. Kota Surabaya memiliki kasus TB terbanyak di Provinsi Jawa Timur yaitu 3.990 kasus. Sedangkan

data Dinas Kesehatan Provinsi Jawa Timur tahun 2013 menunjukkan kasus TB mencapai 41.472 kasus dengan BTA positif baru sebanyak 25.618 kasus dan 1.233 penderita TB meninggal. Salah satu faktor penyebab berkembangnya penyakit ini adalah kurangnya ventilasi dan pencahayaan matahari pada rumah penduduk, serta kurangnya istirahat. Meski terindikasi banyak menyerang warga di pemukiman kumuh, golongan atau strata penduduk ekonomi menengah juga tak luput dari target bakteri ini. TB juga akan menyerang kalangan usia produktif yang bisa mempengaruhi produktifitas kerja penderitanya

Tingginya kasus TB di Jawa Timur menjadi perhatian khusus bagi pemerintah khususnya Dinas Kesehatan Provinsi Jawa Timur. Tindakan preventif yang dapat dilakukan oleh pemerintah adalah dengan mengendalikan faktor-faktor yang mempengaruhi TB. Salah satu metode statistika yang mempelajari hubungan antara variabel dependen dengan variabel independen dalam penelitian ini yaitu jumlah kasus TB di Jawa Timur dengan faktor-faktor yang mempengaruhi jumlah kasus TB di Jawa Timur yaitu analisis regresi. Tujuan dari analisis regresi adalah mengetahui pengaruh perubahan nilai variabel respon yang dipengaruhi oleh perubahan nilai variabel prediktornya (Draper & Smith, 1992).

Penelitian sebelumnya mengenai TB pernah dilakukan Pardeshi (2009) yang meneliti ketahanan hidup penderita TB di India dan Permatasari (2005) yang mengkaji tentang cara pemberantasan TB dengan beberapa strategi. Penelitian mengenai faktor-faktor yang mempengaruhi kasus TB Paru juga dilakukan oleh Faris Muaz (2014) dengan studi kasus puskesmas wilayah Kecamatan Serang Kota Serang. Sedangkan berdasarkan penelitian Nurmala (2009) faktor-faktor Tuberkulosis pada anak adalah umur, status gizi, tingkat pendidikan orang tua yang rendah, tingkat pengetahuan orang tua yang rendah tentang penyakit TB dan kepadatan hunian. Penelitian Amrullah (2011) faktor-faktor resiko penderita TB yaitu pada kelompok produktif (15-50 tahun), jenis kelamin laki-laki, status gizi yang buruk,

diabetes melitus, dan kondisi rumah. Penelitian Soemirat (2000), beberapa faktor penyakit resiko penyakit TB di Amerika yaitu umur, jenis, kelamin, ras, asal negara bagian.

Secara umum, beberapa penelitian yang telah dilakukan tentang TB menyatakan bahwa faktor-faktor yang berpengaruh terhadap terjangkitnya TB pada seseorang adalah faktor lingkungan diantaranya lingkungan fisik meliputi ventilasi, suhu, pencahayaan, dan kelembaban; karakteristik individu meliputi usia, jenis kelamin, kontak penderita, riwayat imunisasi, perilaku, dan status gizi; dan lingkungan sosial meliputi kepadatan penghuni, pendidikan, pengetahuan, dan penghasilan (Sugiarto, 2004). Penelitian selanjutnya oleh Prabu (2008) menyatakan variabel-variabel yang diduga mempengaruhi TB meliputi faktor umur, jenis kelamin, tingkat pendidikan, pekerjaan, kebiasaan merokok, kepadatan hunian kamar, ventilasi, kondisi rumah, kelembaban udara, status gizi, keadaan sosial ekonomi, dan perilaku. Hasil penelitian yang dilakukan oleh Siswanto (2008) menyatakan bahwa meningkatnya penularan infeksi TB Paru yang dilaporkan saat ini, banyak dihubungkan dengan kondisi sosial ekonomi, belum optimalnya fasilitas pelayanan kesehatan masyarakat, meningkatnya jumlah penduduk yang tidak mempunyai tempat tinggal, dan epidemi dari infeksi HIV.

Berdasarkan uraian di atas, peneliti bermaksud melakukan penelitian untuk mengetahui faktor-faktor yang mempengaruhi jumlah kasus TB di Jawa Timur. Faktor-faktor yang menurut peneliti diduga mempengaruhi peningkatan jumlah kasus TB yaitu persentase keluarga dengan kepemilikan sarana sanitasi dasar sehat, persentase penduduk miskin, persentase gizi buruk, persentase tenaga kesehatan terlatih dan persentase rumah tangga ber-PHBS (Perilaku Hidup Bersih dan Sehat). Metode yang digunakan untuk menganalisis faktor-faktor yang mempengaruhi jumlah kasus TB di Jawa Timur adalah metode regresi. Namun karena dalam penelitian ini tidak ada jaminan bahwa hubungan antara variabel prediktor dan variabel respon akan memiliki pola data tertentu, sehingga dalam penelitian ini digunakan regresi

nonparametrik *spline*. Regresi nonparametrik *spline* digunakan untuk mengetahui hubungan variabel respon dan variabel prediktor yang tidak diketahui bentuk pola fungsi regresinya. Metode regresi nonparametrik *spline* dapat mengetahui faktor-faktor yang paling signifikan mempengaruhi terjadinya penyakit TB di Jawa Timur, sehingga diharapkan dengan adanya penelitian ini dapat memberikan informasi kepada pemerintah dalam upaya meminimalisasi terjadinya penyakit TB di Jawa Timur.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan sebelumnya, maka permasalahan pokok yang akan dibahas dalam penelitian ini adalah sebagai berikut.

1. Bagaimana karakteristik dari jumlah kasus TB (Tuberkulosis) di Jawa Timur dan faktor-faktor yang mempengaruhi?
2. Bagaimana pemodelan jumlah kasus TB (Tuberkulosis) di Jawa Timur dan faktor-faktor yang diduga mempengaruhinya dengan menggunakan metode regresi nonparametrik *spline*?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah, tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Mendeskripsikan karakteristik dari jumlah kasus TB (Tuberkulosis) di Jawa Timur dan faktor-faktor yang diduga mempengaruhi.
2. Memodelkan jumlah kasus TB (Tuberkulosis) di Jawa Timur dan faktor-faktor yang diduga mempengaruhi dengan menggunakan metode regresi nonparametrik *spline*.

1.4 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut.

1. Manfaat untuk Pemerintah Jawa Timur

Hasil dari analisis dan pemodelan dalam penelitian ini dapat digunakan sebagai alternatif bagi pemerintah Jawa Timur dalam memutuskan kebijakan dalam rangka menurunkan jumlah kasus TB yang masih terbilang tinggi.

2. Manfaat untuk Bidang Keilmuan

Hasil penelitian ini diharapkan dapat memberikan informasi dan pengetahuan mengenai penerapan metode regresi nonparametrik *spline* untuk memodelkan jumlah kasus TB di Jawa Timur sehingga dapat dijadikan referensi untuk penelitian selanjutnya.

1.5 Batasan Penelitian

Batasan yang digunakan dalam penelitian ini adalah pemilihan titik knot optimal dengan menggunakan metode *Generalized Cross Validation* (GCV) dan banyak knot yang digunakan adalah satu knot, dua knot, tiga knot, dan kombinasi knot.

(halaman ini sengaja dikosongkan)

BAB II

TINJAUAN PUSTAKA

Pada bagian ini akan membahas kajian pustaka yang digunakan untuk menyelesaikan permasalahan pada rumusan masalah penelitian mengenai jumlah kasus tuberkulosis di Jawa Timur meliputi analisis regresi, regresi parametrik, regresi non-parametrik, pemilihan titik knot optimal, uji signifikansi parameter, uji asumsi residual, dan kajian nonstatistik yakni tuberkulosis dan beberapa penelitian sebelumnya yang meneliti tentang tuberkulosis.

2.1 Analisis Regresi

Analisis regresi merupakan sebuah metode statistika yang memberikan penjelasan tentang pola hubungan (model) antara dua variabel atau lebih (Drapper & Smith, 1992). Hubungan antara beberapa variabel tersebut dibentuk dalam model matematis. Pada model ini, terdapat dua variabel yaitu variabel respon dan variabel prediktor. Variabel respon biasa disebut variabel yang bergantung pada variabel prediktor, sedangkan variabel prediktor biasa disebut variabel bebas. Tujuan umum analisis regresi adalah memodelkan variabel respon dengan variabel prediktor dan digunakan untuk memprediksi. Langkah pertama yang harus dilakukan dalam analisis regresi adalah menentukan pola data untuk dapat memutuskan metode analisis yang akan digunakan. Dalam regresi, pola hubungan dua atau lebih variabel prediktor dengan variabel respon tidak selalu berpola parametrik. Terdapat juga beberapa kasus dimana satu atau lebih variabel prediktor memiliki pola nonparametrik. Beberapa kasus lain juga sering ditemui dengan pola data bertipe semiparametrik (Budiantara, 2005). Berikut adalah persamaan regresi secara umum dalam bentuk matriks.

$$Y = X\beta + \varepsilon \quad (2.1)$$

dimana,

Y : vektor kolom berukuran $(n \times 1)$ yang berisikan data respon

X : matrik berukuran $(n \times (q + 1))$

β : nilai estimasi parameter regresi berukuran $((q + 1) \times 1)$

ε : vektor kolom berukuran $(n \times 1)$ berupa *random error*

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1q} \\ 1 & x_{21} & x_{22} & \cdots & x_{2q} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nq} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_q \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Estimasi terhadap parameter $\tilde{\beta}$ dilakukan dengan menggunakan metode kudrat terkecil atau *Ordinary Least Square* (OLS) yaitu dengan tujuan meminimumkan jumlah kuadrat residual. Berdasarkan persamaan 2.1 nilai residual dapat dituliskan dalam persamaan 2.2.

$$\varepsilon = Y - X\beta \quad (2.2)$$

dengan menggunakan metode OLS dapat didefinisikan kuadrat residual dan meminimumkannya seperti persamaan 2.3 sebagai berikut.

$$\begin{aligned} \min_{\beta} (\varepsilon' \varepsilon) &= \min_{\beta} \{ (Y - X\beta)' (Y - X\beta) \} \\ &= Y'Y - Y'X\beta - \beta'X'Y + \beta'X'X\beta \\ &= Y'Y - 2Y'X\beta + \beta'X'X\beta \end{aligned} \quad (2.3)$$

Langkah selanjutnya adalah menurunkan persamaan 2.3 terhadap β sama dengan nol, sehingga didapatkan persamaan 2.4 untuk nilai estimasi parameter β sebagai berikut (Gujarati, 2004 dalam Setiawan, 2010).

$$\begin{aligned} \frac{\partial (\varepsilon' \varepsilon)}{\partial \beta} &= \frac{\partial (Y'Y - 2Y'X\beta + \beta'X'X\beta)}{\partial \beta} = 0 \\ -2Y'X + 2\beta'X'X &= 0 \\ \beta'X'X &= Y'X \\ X'X\beta &= X'Y \\ \hat{\beta} &= (X'X)^{-1} X'Y \end{aligned} \quad (2.4)$$

2.2 Regresi Nonparametrik Spline

Regresi nonparametrik merupakan metode yang dapat digunakan apabila pola data yang diketahui bentuk kurva regresinya atau tidak terdapat informasi masa lalu yang lengkap tentang bentuk pola data (Eubank, 1988). Menurut (Budiantara, 2009), dalam teori estimasi biarkan data mencari sendiri bentuk estimasi dari kurva regresinya tanpa dipengaruhi oleh faktor subyektifitas peneliti. Spline memiliki fleksibilitas yang tinggi, mampu menangani data atau fungsi yang mulus (*smooth*), dan memiliki kemampuan yang sangat baik dalam menangani data yang perilakunya berubah-ubah pada sub interval tertentu. Bentuk model regresi nonparametrik secara umum disajikan sebagai berikut

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.5)$$

dimana y adalah variabel respon dan $f(x_i)$ adalah fungsi spline berorde m dengan titik knot k_1, k_2, \dots, k_r yang diberikan oleh persamaan

$$f(x_i) = \sum_{j=0}^m \beta_j x_i^j + \sum_{k=1}^r \beta_{m+k} (x_i - k_k)_+^m. \quad (2.6)$$

Apabila Persamaan 2.6 disubstitusikan ke dalam Persamaan 2.5 maka akan diperoleh model regresi nonparametrik spline sebagai berikut

$$y_i = \sum_{j=0}^m \beta_j x_i^j + \sum_{k=1}^r \beta_{m+k} (x_i - k_k)_+^m + \varepsilon_i, \quad i = 1, 2, \dots, n. \quad (2.7)$$

Fungsi $(x_i - k_k)_+^m$ merupakan fungsi potongan yang diberikan oleh

$$(x_i - k_k)_+^m = \begin{cases} (x_i - k_k)^m, & x_i \geq k_k \\ 0, & x_i < k_k \end{cases}. \quad (2.8)$$

2.3 Pemilihan Titik Knot Optimal

Metode yang baik untuk memilih titik knot optimal yaitu *Generalized Cross Validation* (GCV). Titik knot merupakan titik perpaduan bersama dimana terdapat perubahan perilaku pola kurva pada internal yang berlainan (Budiantara, 2006). Model spline terbaik dengan titik knot optimal adalah ketika GCV minimum. Adapun fungsi GCV diberikan dalam Persamaan 2.9 sebagai berikut

$$GCV(k_1, k_2, \dots, k_r) = \frac{MSE(k_1, k_2, \dots, k_r)}{(n^{-1} \text{tr} [I - A(k_1, k_2, \dots, k_r)])^2} \quad (2.9)$$

Berdasarkan Persamaan 2.9 titik (k_1, k_2, \dots, k_r) merupakan titik knot sedangkan matriks $A(k_1, k_2, \dots, k_r)$ diperoleh dari persamaan berikut

$$\hat{y} = A(k_1, k_2, \dots, k_r) y. \quad (2.10)$$

Nilai dari $MSE(k_1, k_2, \dots, k_r)$ adalah sebagai berikut

$$MSE(k_1, k_2, \dots, k_r) = n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad (2.11)$$

2.4 Pengujian Estimasi Parameter

Uji parameter model dilakukan untuk menentukan variabel prediktor yang berpengaruh terhadap variabel respon. Pada regresi spline, uji parameter model dilakukan setelah mendapatkan model regresi dengan titik knot optimal berdasarkan nilai GCV yang paling minimum. Terdapat dua tahap pengujian parameter yaitu pengujian secara serentak dan secara parsial.

2.4.1 Pengujian Serentak (Simultan)

Uji serentak digunakan untuk mengetahui apakah parameter model regresi signifikan atau tidak. Hipotesis yang digunakan dalam pengujian ini adalah sebagai berikut

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{m+r} = 0$$

H_1 : minimal ada satu $\beta_j \neq 0, j = 1, 2, \dots, m + r$

dimana nilai $m + r$ adalah banyak parameter dalam model regresi dan statistik uji yang digunakan adalah uji F sebagai berikut

$$F_{hitung} = \frac{MS_{Reg}}{MS_{Error}}. \quad (2.12)$$

H_0 akan ditolak jika $F_{hitung} > F_{\alpha; ((m+r), n-(m+r)-1)}$. *Analysis of variance* (ANOVA) dari model regresi disajikan dalam Tabel 2.1.

Tabel 2.1 *Analysis of Variance*

Sumber Variasi	Df	Sum of Square	Mean Square	F hitung
Regresi	$m + r$	$\hat{\beta}'X'y - n\bar{y}^2$	$\frac{SSR}{db_{regresi}}$	$\frac{MS_{Reg}}{MS_{Error}}$
Error	$n - (m + r) - 1$	$y'y - \hat{\beta}'X'y$	$\frac{SSE}{db_{error}}$	—
Total	$n - 1$	$y'y - n\bar{y}^2$	—	—

2.4.2 Pengujian Individu

Uji parsial digunakan untuk mengetahui variabel yang memberikan pengaruh signifikan terhadap model. Hipotesis yang digunakan dalam pengujian ini adalah sebagai berikut

$H_0 : \beta_j = 0$

$H_1 : \beta_j \neq 0, j = 1, 2, \dots, m + r$.

Statistik uji:

$$t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (2.13)$$

dimana $\hat{\beta}_j$ adalah penaksir untuk parameter β_j dan $SE(\hat{\beta}_j)$ adalah akar dari elemen diagonal ke- j dari matriks $\sigma^2(X'X)^{-1}$. H_0 ditolak jika $|t_{hitung}| > t_{\frac{\alpha}{2}; (n-(m+r)-1)}$.

2.5 Kriteria Kebaikan Model

Salah satu ukuran yang sering digunakan untuk mengetahui kebaikan model yaitu koefisien determinasi atau R^2 . Koefisien determinasi digunakan untuk mengetahui sampai sejauh mana kecocokan regresi yang terbentuk terhadap data observasi. Koefisien determinasi menggambarkan bagian dari variasi total yang dapat diterangkan model. Semakin tinggi nilai R^2 yang dihasilkan suatu model, maka semakin baik pula variabel-variabel prediktor dalam model tersebut dalam menjelaskan variabilitas variabel respon (Drapper dan Smith, 1992). Berikut ini merupakan rumus untuk menghitung R^2

$$R^2 = \frac{SS_{regresi}}{SS_{total}} = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} \quad (2.14)$$

2.6 Pemeriksaan Asumsi Residual Dalam Model Regresi

Residual menentukan seberapa baik model tersebut menjelaskan hubungan antara variabel respon dan variabel prediktor (McClave, et al., 2010). Pengujian asumsi residual (*goodness of fit*) dilakukan untuk mengetahui apakah residual yang dihasilkan telah memenuhi asumsi yakni identik, independen, dan berdistribusi normal (IIDN).

2.6.1 Asumsi Residual Identik

Salah satu asumsi regresi yang harus dipenuhi adalah homogenitas variansi dari *error* (homoskedastisitas). Homoskedastisitas berarti bahwa variansi dari *error* bersifat konstan (tetap) atau disebut juga identik. Salah satu cara untuk mendeteksi adanya heteroskedastisitas adalah dengan menggunakan uji Glejser. Uji Glejser secara umum dilakukan dengan meregresikan nilai mutlak error dengan variabel prediktor (Gujarati, 2004). Hipotesis yang digunakan adalah sebagai berikut

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 : \text{minimal terdapat satu } \sigma_i^2 \neq \sigma^2$$

Nilai statistik uji Glejser dapat dirumuskan sebagai berikut

$$F_{hitung} = \frac{\left[\sum_{i=1}^n (|\hat{e}_i| - |\bar{e}|)^2 \right] / ((m+r)-1)}{\left[\sum_{i=1}^n (|e_i| - |\hat{e}_i|)^2 \right] / (n-(m+r))} \quad (2.15)$$

dimana n adalah banyak observasi dan $(m+r)$ adalah banyak parameter model Glejser dengan daerah penolakan yaitu H_0 ditolak jika $F_{hitung} > F_{tabel}(F_{\alpha, ((m+r)-1, n-(m+r))})$.

2.6.2 Asumsi Residual Independen

Asumsi selanjutnya yang harus dipenuhi adalah tidak terdapat korelasi antar residual atau autokorelasi. Pendeteksian autokorelasi dapat dilakukan dengan membuat plot *Autocorrelation Function* (ACF). Apabila ada nilai autokorelasi yang keluar dari batas signifikansi maka dapat dikatakan terdapat kasus autokorelasi. Sebaliknya jika tidak terdapat nilai autokorelasi yang keluar batas signifikansi maka tidak terdapat kasus autokorelasi. Interval konfidensi $(1-\alpha)$ 100% untuk autokorelasi ρ_s diberikan oleh

$$-t_{n-1; \frac{\alpha}{2}} SE(\hat{\rho}_s) < \rho_s < t_{n-1; \frac{\alpha}{2}} SE(\hat{\rho}_s). \quad (2.16)$$

2.6.3 Asumsi Residual Berdistribusi Normal

Residual dari suatu model regresi harus mengikuti distribusi normal dengan nilai mean nol dan varians σ^2 . Pengujian distribusi normal ini dapat dilakukan secara visual dengan *Normal Probability Plot Residual*. Residual dari model dikatakan telah memenuhi asumsi distribusi normal apabila plot terlihat cenderung mengikuti garis lurus 45° . Cara lain yang sering digunakan adalah pengujian distribusi normal Kolmogorov-Smirnov (Daniel, 1990). Berikut ini merupakan hipotesis untuk Uji Kolmogorov-Smirnov.

$H_0 : F_0(x) = F(x)$ (Residual berdistribusi Normal)

$H_1 : F_0(x) \neq F(x)$ (Residual tidak berdistribusi Normal)

Sedangkan statistik uji yang digunakan untuk Uji Kolmogorov-Smirnov adalah

$$D = \text{Sup} |S(x) - F_0(x)|. \quad (2.17)$$

Dimana

D : Jarak vertikal terjauh antara $S(x)$ dan $F_0(x)$

$F_0(x)$: Fungsi peluang kumulatif berdistribusi normal

$S(x)$: Fungsi peluang kumulatif yang dihitung dari data sampel

Tolak H_0 adalah apabila $|D| > q_{(1-\alpha)}$ dengan nilai $q_{(1-\alpha)}$ didapatkan dari tabel Kolmogorov-Smirnov.

2.7 Tuberkulosis (TB)

Tuberkulosis (TB) adalah penyakit radang parenkim paru yang menular karena infeksi kuman TB yaitu *Mycrobacterium tuberculosis*. Sebagian besar kuman TB menyerang paru, tetapi dapat juga mengenai organ tubuh lainnya. Keluhan yang dirasakan penderita TB dapat bermacam-macam, seperti demam, batuk, sesak napas, nyeri dada, dan malaise. Sumber penularan adalah pasien TB BTA (Basil Tahan Asam). Pada waktu batuk atau bersin, pasien menyebarkan kuman ke udara dalam bentuk percikan dahak. Sekali batuk dapat menghasilkan sekitar 3.000 percikan dahak. Kasus TB terutama terjadi pada usia produktif kerja, yaitu kelompok umur 15-55 tahun yang berdampak pada SDM sehingga dapat mengganggu perekonomian keluarga, masyarakat, dan negara. Berdasarkan laporan WHO (2010), terdapat 22 negara dengan kategori beban tinggi terhadap tuberkulosis. Indonesia sebagai salah satu negara yang masuk dalam kategori negara beban tinggi terhadap tuberkulosis, berada pada peringkat kelima setelah India, Cina, Afrika Selatan, dan Nigeria dengan jumlah penderita tuberkulosis sebesar 429 orang (Kementrian Kesehatan RI, 2011). Jawa Timur merupakan salah satu provinsi di Indonesia merupakan salah satu penyumbang jumlah penderita tuberkulosis terbanyak kedua dibawah Provinsi Jawa Barat. Jumlah kasus penyakit TB di Jawa Timur sebanyak 41.472 kasus dengan BTA positif baru sebanyak 25.618 kasus

dan 1.233 penderita tuberkulosis meninggal (Dinas Kesehatan Provinsi Jawa Timur, 2013).

Menurut Departemen Kesehatan RI (2009), tuberkulosis dibedakan menjadi dua klasifikasi, yaitu tuberkulosis paru dan tuberkulosis ekstra paru. TB paru adalah penyakit tuberkulosis yang menyerang jaringan paru, sedangkan tuberkulosis ekstra paru merupakan penyakit tuberkulosis yang menyerang organ tubuh lain selain paru, diantaranya organ selaput otak, selaput jantung, (pericardium), kelenjar getah bening, tulang, limfa, persendian, kulit, asus, ginjal, saluran kencing, dan lain-lain. Tuberkulosis dapat menyebabkan kematian apabila tidak segera diobati, dimana 50 persen dari pasien penderita tuberkulosis akan meninggal setelah 5 tahun.

2.8 Penelitian Sebelumnya

Sampai saat ini penyakit TB merupakan masalah yang cukup serius karena menyebabkan kematian yang tinggi pada usia produktif. Penelitian tentang tuberkulosis di Indonesia beberapa telah dilakukan diantaranya adalah penelitian yang dilakukan oleh Hasyim (2010), tentang pemodelan angka kejadian penyakit infeksi tuberkulosis paru di Kabupaten Sorong, Papua Barat dengan menggunakan metode *Multivariate Adaptive Regression Spline* (MARS). Hasil penelitian menunjukkan bahwa jenis pekerjaan, umur, kebiasaan merokok, status sosial ekonomi, konsumsi alkohol dan tingkat pendidikan berpengaruh signifikan terhadap penyakit tuberkulosis paru di Kabupaten Sorong, Provinsi Papua Barat. Selain itu, Dwikentarti (2010) juga menganalisis faktor-faktor yang mempengaruhi penyakit tuberkulosis pada pasien dengan regresi logistik multinomial di Kota Semarang. Dari penelitian diperoleh hasil bahwa variabel umur, tempat tinggal, dan kebiasaan merokok berpengaruh signifikan terhadap BTA positif sebesar 78,41 persen. Penelitian tentang faktor kesehatan lingkungan yang berhubungan dengan kejadian TB Paru di Kabupaten Cilacap dilakukan oleh Fatimah (2008), memberikan hasil bahwa pencahayaan, kelembapan, ventilasi, dan status gizi memiliki hubungan dengan terjadinya

penyakit tuberkulosis. Penelitian terkait TB selanjutnya dilakukan oleh Lestari (2014) yang meneliti faktor-faktor yang mempengaruhi jumlah kasus TB dengan menggunakan metode *Generalized Poisson Regression* (GPR) dan *Geographically Weighted Poisson Regression* (GWPR). Hasil penelitian tersebut menunjukkan bahwa variabel persentase penduduk usia produktif, persentase tenaga kesehatan terdidik, dan persentase TUPM sehat yang berpengaruh signifikan terhadap jumlah kasus TB di Jawa Timur. Penelitian serupa juga pernah dilakukan oleh Sukmawati (2010) dengan mengambil studi kasus di RSUD Ibnu Sina Kabupaten Gresik dan menggunakan metode analisis survival dengan model *Proportional Hazard*. Penelitian tersebut memberikan kesimpulan bahwa pasien yang menderita penyakit TB sebagian besar berjenis kelamin perempuan, sehingga banyak yang memiliki kebiasaan tidak merokok, pendidikan terakhirnya adalah SMP dan memiliki pendapatan keluarga sebesar Rp 1.500.000 sampai dengan Rp 2.500.000. Selain itu sebagian besar pasien memiliki pencahayan dalam rumah, sanitasi, serta kebersihan lingkungan yang belum baik. Sedangkan Puspita (2014), melakukan penelitian yang serupa menggunakan metode regresi logistik biner dengan mengambil studi kasus di RSUD Haji Surabaya. Penelitian tersebut menunjukkan bahwa mayoritas 54,8% pasien didiagnosa TB Paru, berusia dibawah 45 tahun dan berjenis kelamin laki-laki. Sedangkan faktor-faktor risiko yang mempengaruhi adalah usia, kekurangan nutrisi, dan sesak nafas.

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan sekunder yang diperoleh dari Dinas Kesehatan Provinsi Jawa Timur tahun 2014 dan Badan Pusat Statistik Provinsi Jawa Timur tahun 2014. Data tersebut diambil dari 38 kabupaten/kota yang ada di Provinsi Jawa Timur. Data yang digunakan mencakup data mengenai jumlah kasus penderita tuberkulosis dan faktor-faktor yang diduga mempengaruhi tingginya jumlah kasus penderita tuberkulosis di Provinsi Jawa Timur. Variabel-variabel yang digunakan dalam penelitian ini meliputi jumlah kasus penyakit tuberkulosis di Jawa Timur, variabel persentase kegiatan sarana sanitasi dasar keluarga, persentase gizi buruk masyarakat, persentase tenaga kesehatan terlatih TB, dan persentase rumah tangga ber-PHBS yang bersumber dari Dinas Kesehatan Provinsi Jawa Timur tahun 2014. Sementara untuk variabel persentase penduduk miskin di setiap kabupaten/kota di Jawa Timur berasal dari Badan Pusat Statistik Jawa Timur tahun 2014.

3.2 Variabel Penelitian

Variabel respon yang digunakan dalam penelitian ini adalah jumlah kasus penyakit tuberkulosis (Y) di Provinsi Jawa Timur dengan 5 variabel yang prediktor yang diduga mempengaruhi jumlah kasus penyakit tuberkulosis di Jawa Timur. Variabel prediktor yang digunakan menggunakan satuan persen, sedangkan untuk variabel respon menggunakan satuan jumlah. Variabel-variabel yang digunakan dalam penelitian ini tercantum Tabel 3.1.

Tabel 3.1 Variabel Penelitian

Variabel	Nama Variabel	Skala Pengukuran
Y	Jumlah Kasus Baru Penyakit Tuberkulosis	Rasio
X_1	Persentase Keluarga dengan Kepemilikan Sarana Sanitasi Dasar Sehat	Rasio
X_2	Persentase Penduduk Miskin	Rasio
X_3	Persentase Gizi Buruk Masyarakat	Rasio
X_4	Persentase Tenaga Kesehatan Terlatih	Rasio
X_5	Persentase Rumah Tangga ber-PHBS	Rasio

Selanjutnya struktur data yang digunakan dalam penelitian disajikan pada Tabel 3.2.

Tabel 3.2 Struktur Data Variabel Penelitian

Kab/Kota	Y	X_1	X_2	X_3	...	X_5
1	y_1	$x_{1(1)}$	$x_{2(1)}$	$x_{3(1)}$...	$x_{5(1)}$
2	y_2	$x_{1(2)}$	$x_{2(2)}$	$x_{3(2)}$...	$x_{5(2)}$
3	y_3	$x_{1(3)}$	$x_{2(3)}$	$x_{3(3)}$...	$x_{5(3)}$
4	y_4	$x_{1(4)}$	$x_{2(4)}$	$x_{3(4)}$...	$x_{5(4)}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
38	y_{38}	$x_{1(38)}$	$x_{2(38)}$	$x_{3(38)}$...	$x_{5(38)}$

3.3 Definisi Operasional Variabel

Deskripsi variabel yang akan digunakan dalam penelitian ini adalah sebagai berikut.

- Variabel Y menyatakan jumlah kasus baru penyakit TB Paru. TB Paru kasus baru merupakan pasien yang belum pernah berobat menggunakan obat anti tuberkulosis (OAT) atau pernah menggunakan OAT dalam jangka waktu kurang dari 1 bulan.
- Variabel X_1 menyatakan persentase keluarga dengan kepemilikan sarana sanitasi dasar sehat. Menurut Dinas Kesehatan Provinsi Jawa Timur (2013), keluarga dengan kepemilikan sarana sanitasi dasar meliputi kepemilikan jamban sehat, tempat sampah dan pengolahan air limbah

keluarga. Pengolaan air limbah merupakan salah satu komponen utama bagi kesehatan lingkungan. Kurangnya kesadaran masyarakat terkait kesehatan lingkungan dan sanitasi dasar menyebabkan tingginya jumlah penyakit di masyarakat khususnya penyakit TB yang terus mengalami peningkatan tiap tahun.

- c. Variabel X_2 menyatakan persentase penduduk miskin yaitu, persentase penduduk yang memiliki ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran. Jika penduduk memiliki rata-rata pengeluaran perkapita pengeluaran di bawah garis kemiskinan tergolong dalam penduduk miskin.
- d. Variabel X_3 menyatakan persentase gizi buruk masyarakat. Status gizi balita digunakan untuk merepresentasikan status gizi masyarakat karena kelompok balita merupakan kelompok yang mudah mengalami perubahan status gizi dibandingkan kelompok umur lain. Status gizi balita diukur menggunakan indikator berat badan menurut tinggi badan dengan standar yang telah ditentukan. Status gizi buruk terjadi apabila berat badan berdasarkan tinggi badan jauh di bawah standar yang telah ditentukan (Direktorat Bina, 2013). Semakin buruk status gizi seseorang, maka akan semakin mudah orang tersebut terserang penyakit, terutama penyakit TB.
- e. Variabel X_4 menyatakan persentase tenaga kesehatan terlatih. Tenaga kesehatan merupakan faktor penggerak utama dalam pencapaian tujuan dan keberhasilan program pembangunan kesehatan (Dinas Kesehatan Provinsi Jawa Timur, 2013). Pelaksanaan pelatihan terhadap tenaga kesehatan bertujuan agar tenaga kesehatan memiliki kompetensi dan paham tentang penyakit TB. Sehingga strategi pengendalian TB dapat tercapai.
- f. Variabel X_5 menyatakan persentase rumah tangga ber-PHBS. Perilaku Hidup Bersih dan Sehat (PHBS) adalah semua perilaku kesehatan yang dilakukan atas kesadaran semua

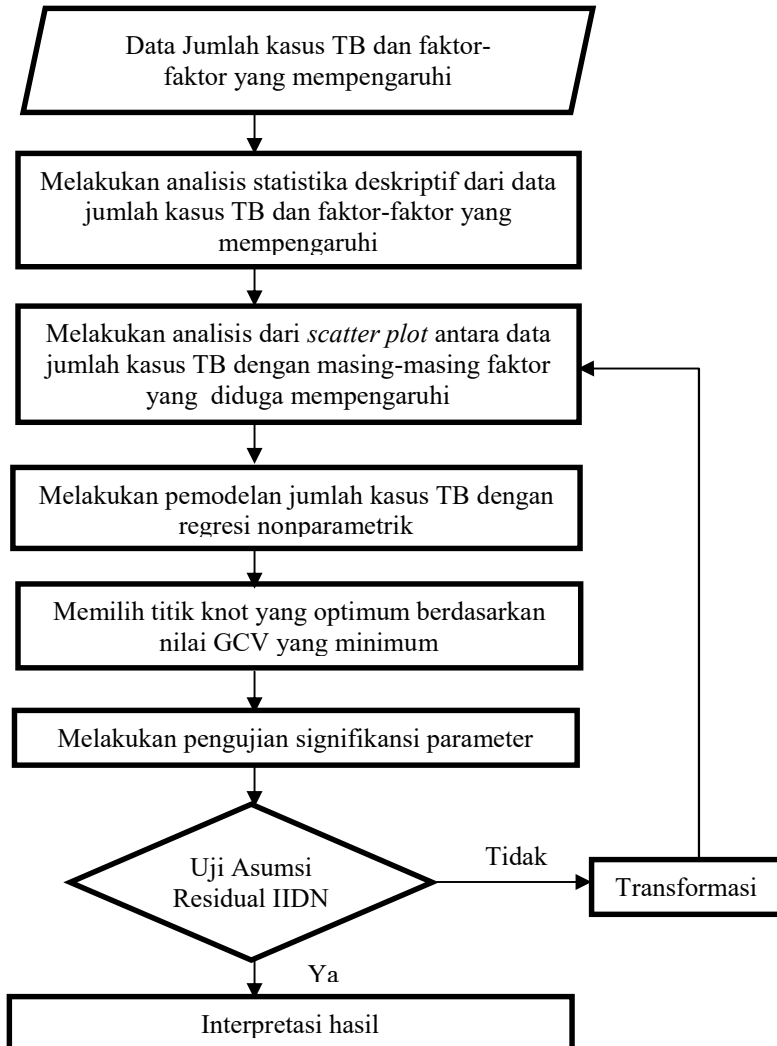
anggota keluarga dan masyarakat. PHBS sangat diperlukan karena dapat mencegah dan menaggulangi penyakit. Persentase rumah tangga yang ber-PHBS diperoleh dari perbandingan antara jumlah rumah tangga yang melaksanakan 10 indikator PHBS dibagi dengan jumlah rumah tangga yang diperiksa.

3.4 Langkah Analisis dan Diagram Alir

Langkah-langkah analisis yang dilakukan pada penelitian ini adalah sebagai berikut:

1. Melakukan analisis statistika deskriptif untuk mengetahui karakteristik data meliputi rata-rata, varians, nilai maksimum dan nilai minimum pada variabel jumlah kasus TB di Jawa Timur dan faktor yang diduga mempengaruhinya.
2. Membuat *scatterplot* antara jumlah kasus TB di Jawa Timur dengan masing-masing variabel prediktor yang diduga mempengaruhi untuk mengetahui pola hubungan data antara variabel respon dengan masing-masing variabel prediktornya.
3. Melakukan pemodelan jumlah kasus TB dengan regresi nonparametrik menggunakan pendekatan *spline* dengan satu, dua, tiga, dan kombinasi knot.
4. Memilih titik knot optimal dengan menggunakan karakteristik nilai GCV paling minimum.
5. Melakukan pemodelan dengan menggunakan titik knot optimal.
6. Menguji signifikansi parameter dari model yang terbentuk.
7. Menguji asumsi residual dari model yang terbentuk.

Tahapan dari metode analisis data selengkapnya disajikan pada Gambar 3.1.



Gambar 3.1 Diagram Alir Analisis Data

(halaman ini sengaja dikosongkan)

BAB IV ANALISIS DAN PEMBAHASAN

Bab ini membahas tentang karakteristik jumlah kasus TB di Jawa Timur tahun 2014 serta faktor-faktor yang diduga berpengaruh menggunakan statistika deskriptif. Selain itu, bab ini juga membahas mengenai pemodelan jumlah kasus TB menggunakan regresi nonparametrik *spline*, dimana kurva regresi nonparametrik diperoleh menggunakan fungsi *spline* linear satu, dua, tiga knot, dan kombinasi knot.

4.1 Karakteristik Jumlah Kasus Tuberkulosis di Jawa Timur

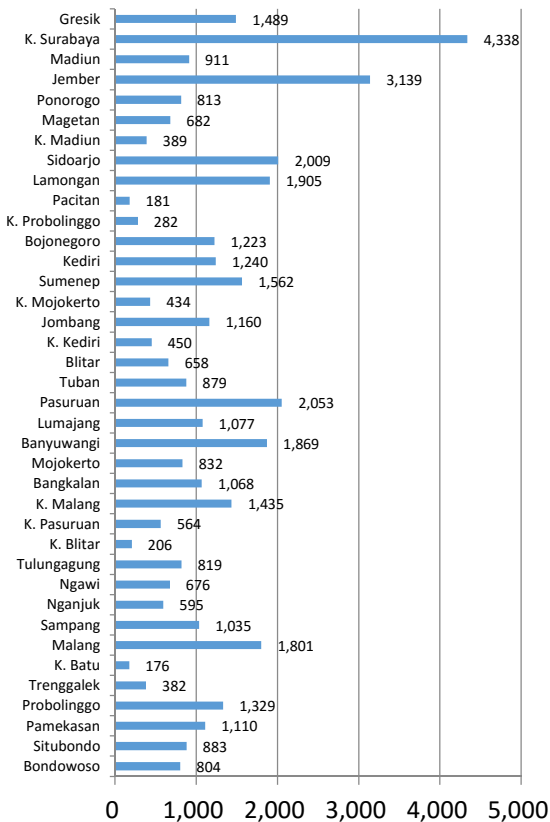
Jawa Timur merupakan salah satu provinsi di Indonesia. Jawa Timur dengan ibukota provinsi Surabaya, terletak di bagian timur pulau Jawa dengan luas wilayah daratan sebesar 47.959 km². Jumlah penderita TB di Provinsi Jawa Timur tahun 2014 sebesar 42.458 kasus dengan kasus BTA positif baru sebanyak 22.866 kasus. Terdapat beberapa faktor yang diduga mempengaruhi tingginya jumlah kasus TB di Jawa Timur. Pada penelitian kali ini, terdapat lima faktor yang diduga mempengaruhi jumlah kasus TB di Jawa Timur. Karakteristik lima faktor tersebut disajikan sebagai berikut.

Tabel 4.1 Karakteristik Jumlah Kasus Tuberkulosis dan Faktor-Faktor yang Mempengaruhi

Variabel	Mean	Varians	Minimum	Maksimum
Y	1.117	682.921	176	4.338
X_1	72,81	235,45	33,70	95,55
X_2	12,096	24,926	4,590	25,800
X_3	2,019	0,915	0,000	4,097
X_4	8,473	10,819	2,721	16,344
X_5	46,29	209,51	20,06	68,67

Berdasarkan Tabel 4.1 variabel Y merupakan variabel jumlah kasus tuberkulosis di Jawa Timur dimana rata-rata jumlah kasus TB di Jawa Timur tahun 2014 sebesar 1.117 kasus di setiap

kabupaten/kota dengan variasi yang sangat besar artinya, di daerah tertentu jumlah kasusnya ada yang sangat banyak dan di daerah lainnya jumlah kasusnya sedikit maka dapat diduga bahwa kasus TB mewabah pada suatu kabupaten/kota tertentu di Jawa Timur. Hal ini dapat dilihat pada nilai maksimum jumlah kasus TB di Jawa Timur yaitu kota Surabaya sedangkan nilai minimum jumlah kasus TB di Jawa Timur yaitu kota Batu. Secara visual jumlah kasus TB di Jawa Timur tahun 2014 disajikan dalam diagram batang pada Gambar 4.1 yang selengkapnya dapat dilihat pada Lampiran 1.



Gambar 4.1 Jumlah Kasus Tuberkulosis di Jawa Timur

Gambar 4.1 menampilkan diagram batang jumlah kasus tuberkulosis di Jawa Timur pada setiap kabupaten/kota tahun 2014. Kota Surabaya merupakan kota yang memiliki jumlah kasus tuberkulosis terbanyak yaitu sebesar 4.338 kasus. Tingginya jumlah kasus tuberkulosis di dapat disebabkan karena Kota Surabaya merupakan salah satu kota dengan tingkat kepadatan penduduk tertinggi di Jawa Timur yaitu sebesar 8683 jiwa/km². Sedangkan kota dengan jumlah kasus tuberkulosis terendah adalah Kota Batu sebanyak 176 kasus.

Pada Tabel 4.1 dapat dijelaskan bahwa persentase keluarga dengan kepemilikan sarana sanitasi dasar sehat (X_1) tertinggi terdapat di Kabupaten Magetan yaitu sebesar 95,55% sedangkan persentase keluarga dengan kepemilikan sarana sanitasi dasar sehat terendah adalah di Kabupaten Situbondo yaitu sebesar 33,7%. Provinsi dengan persentase penduduk miskin (X_2) tertinggi sebesar 25,8% yaitu pada Kabupaten Sampang dan terendah sebesar 4,59% yaitu pada Kota Batu. Persentase gizi buruk masyarakat (X_3) paling tinggi adalah Kota Pasuruan dengan sekitar 4,10% namun berbeda dengan Kabupaten Bojonegoro yang dimana persentase gizi buruk masyarakatnya adalah 0,00%, artinya status gizi masyarakat di Kabupaten Bojonegoro terbilang baik. Untuk persentase tenaga kesehatan terlatih TB (X_4), persentase tertinggi adalah di Kabupaten Sumenep yaitu sebesar 16,34% dan terendah adalah Kabupaten Sidoarjo sebesar 2,72%. Sedangkan pada persentase rumah tangga ber-PHBS (X_5), kabupaten dengan persentase rumah tangga ber-PHBS terbanyak adalah Kabupaten Gresik dan kabupaten dengan persentase rumah tangga ber-PHBS terendah adalah Kabupaten Bondowoso.

Berdasarkan Tabel 4.1 diketahui bahwa rata-rata persentase keluarga dengan kepemilikan sarana sanitasi dasar sehat adalah 72,81% dengan varians yang cukup besar yaitu 235,45 yang menandakan keragaman data yang tinggi. Rata-rata persentase penduduk miskin di Jawa Timur adalah 12,09% dengan varians sebesar 24,926. Status gizi masyarakat dapat memberikan gambaran terhadap derajat kesehatan masyarakat di suatu

wilayah. Rata-rata persentase gizi buruk masyarakat di Jawa Timur adalah 2,01% di setiap kabupaten/kota dengan varians yaitu 0,91% yang menunjukkan bahwa keragaman data tidak tinggi. Variabel lain yang diduga mempengaruhi tingginya jumlah kasus tuberkulosis di Jawa Timur adalah tenaga kesehatan terlatih TB. Tenaga kesehatan merupakan faktor penggerak utama dalam pencapaian tujuan dan keberhasilan program pembangunan kesehatan. Diketahui bahwa rata-rata persentase tenaga kesehatan terlatih TB adalah 8,47% dengan varians 10,81. Selain itu, jumlah rumah tangga ber-PHBS di setiap kabupaten/kota juga diduga mempengaruhi jumlah kasus tuberkulosis di Jawa Timur. Persentase rumah tangga yang ber-PHBS diperoleh dari perbandingan antara jumlah rumah tangga yang melaksanakan 10 indikator PHBS dibagi dengan jumlah rumah tangga yang diperiksa. Rata-rata persentase rumah tangga ber-PHBS di setiap kabupaten/kota di Jawa Timur adalah 46,29% dengan varians yang cukup besar yaitu 209,51.

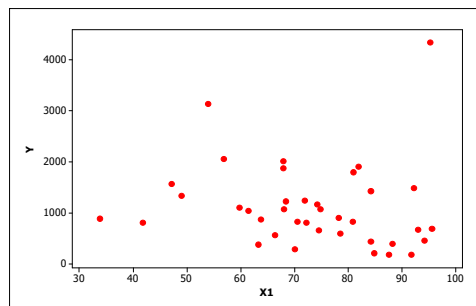
Untuk setiap faktor yang diduga mempengaruhi jumlah kasus tuberkulosis di Jawa Timur, seharusnya sesuai dengan teori, seperti: semakin tinggi persentase keluarga dengan kepemilikan sarana sanitasi dasar sehat maka jumlah kasus tuberkulosis semakin turun, semakin rendah persentase penduduk miskin maka jumlah kasus tuberkulosis semakin menurun, semakin rendah persentase gizi buruk masyarakat maka jumlah kasus tuberkulosis semakin turun, semakin tinggi persentase tenaga kesehatan terlatih TB maka jumlah kasus tuberkulosis semakin turun, serta semakin tinggi persentase rumah tangga ber-PHBS maka jumlah kasus tuberkulosis semakin turun. Namun kenyataannya tidak semua kabupaten/kota di Jawa Timur memenuhi teori tersebut. Semua berpola acak dan hanya pada beberapa wilayah saja yang bisa dikatakan bahwa teori tersebut berlaku untuk faktor-faktor yang mempengaruhi jumlah kasus tuberkulosis di Jawa Timur. Oleh karena itu, metode statistik perlu diterapkan untuk mengetahui faktor-faktor yang menyebabkan tingginya jumlah

kasus tuberkulosis di Jawa Timur. Metode yang akan diterapkan adalah regresi nonparametrik spline.

4.2 Analisis Faktor-Faktor yang Diduga Mempengaruhi Jumlah Kasus Tuberkulosis di Jawa Timur

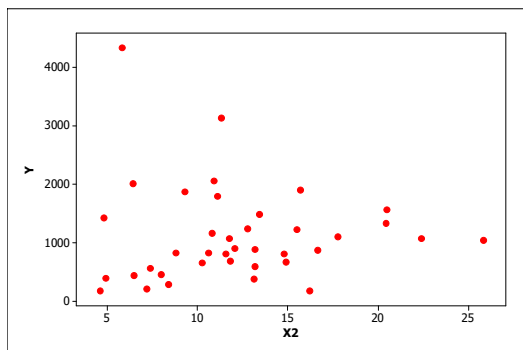
Terdapat beberapa langkah dalam memodelkan faktor-faktor yang mempengaruhi jumlah kasus TB di Jawa Timur yaitu membuat pola data antara jumlah kasus TB di Jawa Timur dengan variabel-variabel yang diduga mempengaruhi, pemilihan titik knot optimal dengan satu titik knot, dua titik knot, tiga titik knot dan kombinasi titik knot. Lalu, akan dilakukan pemilihan model terbaik berdasarkan nilai GCV terkecil, pengujian signifikansi parameter model secara serentak dan individu, serta pengujian asumsi residual.

Berdasarkan hasil *scatterplot* dapat diketahui apakah terbentuk suatu pola tertentu atau tidak. Jika pada *scatterplot* terbentuk suatu pola tertentu seperti linier, kuadratik, kubik atau pola lainnya, maka dalam pemodelan digunakan pendekatan regresi parametrik. Namun, apabila *scatterplot* tidak membentuk pola tertentu, maka digunakan pendekatan regresi nonparametrik. Apabila dalam *scatterplot* terdapat sebagian membentuk pola tertentu dan sebagian sisanya tidak membentuk pola apapun, maka digunakan pendekatan regresi semiparametrik.



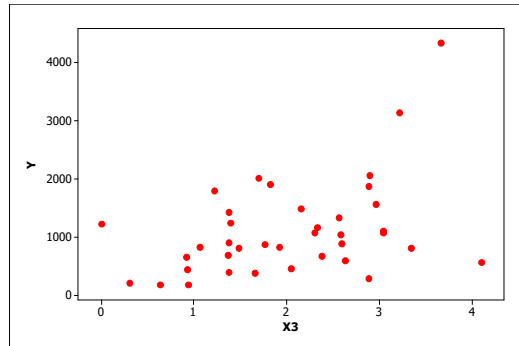
Gambar 4.2 *Scatterplot* Persentase Keluarga Dengan Kepemilikan Sarana Sanitasi Dasar Sehat (X_1)

Gambar 4.2 menyajikan *scatterplot* yang menunjukkan pola hubungan antara jumlah kasus TB di Jawa Timur (Y) dengan variabel prediktor persentase keluarga dengan kepemilikan sarana sanitasi dasar sehat (X_1). Berdasarkan Gambar 4.2 diketahui bahwa antara variabel respon dengan variabel prediktor tidak membentuk suatu pola tertentu. Karena tidak diketahui bentuk pola datanya, sehingga estimasi model tidak dilakukan menggunakan pendekatan regresi parametrik. Metode yang digunakan untuk melakukan estimasi model terhadap variabel X_1 adalah menggunakan pendekatan regresi nonparametrik. Berikut ini merupakan Gambar 4.3 menunjukkan *scatterplot* antara variabel respon jumlah jumlah kasus TB dengan variabel prediktor berupa persentase penduduk miskin (X_2).



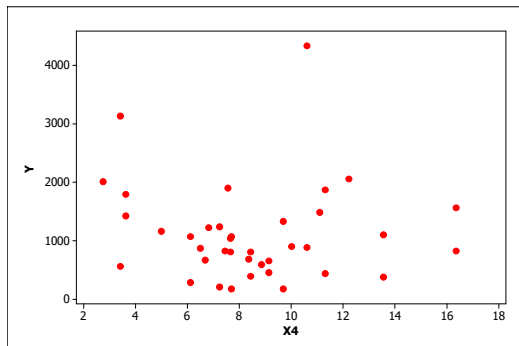
Gambar 4.3 *Scatterplot* Persentase Penduduk Miskin (X_2)

Berdasarkan Gambar 4.3 diketahui bahwa antara variabel respon jumlah kasus TB di Jawa Timur dengan variabel prediktor persentase penduduk miskin (X_2) tidak membentuk suatu pola tertentu, sehingga akan dilakukan estimasi model menggunakan regresi nonparametrik.



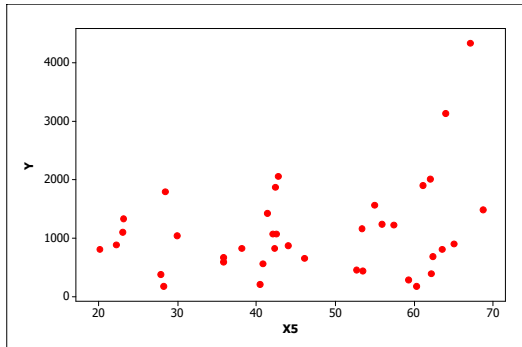
Gambar 4.4 *Scatterplot* Persentase Gizi Buruk Masyarakat (X_3)

Berdasarkan Gambar 4.4 diketahui bahwa antara variabel respon jumlah kasus TB dengan variabel prediktor persentase gizi buruk masyarakat (X_3) memberikan pola data yang tidak membentuk suatu pola tertentu, maka estimasi model menggunakan regresi nonparametrik.



Gambar 4.5 *Scatterplot* Persentase Tenaga Kesehatan Terlatih TB (X_4)

Berdasarkan Gambar 4.5 diketahui bahwa antara variabel respon jumlah kasus TB dengan variabel prediktor persentase tenaga kesehatan terlatih TB (X_4) memberikan pola data yang tidak membentuk suatu pola tertentu, sehingga akan dilakukan estimasi model menggunakan regresi nonparametrik.



Gambar 4.6 Scatterplot Persentase Rumah Tangga Ber-PHBS (X_5)

Berdasarkan Gambar 4.6 diketahui bahwa antara variabel respon jumlah kasus TB dengan variabel prediktor persentase rumah tangga ber-PHBS (X_5) memberikan pola data yang tidak membentuk suatu pola tertentu, maka estimasi model menggunakan regresi nonparametrik adalah metode terbaik yang dapat digunakan.

4.3 Pemilihan Titik Knot Optimum

Titik knot merupakan titik perpaduan dimana terjadi perubahan pola data. Model regresi nonparametrik spline terbaik didapatkan dari titik knot yang optimal. Untuk menentukan titik knot yang optimal, digunakan metode GCV. Berikut ini merupakan pemilihan titik knot optimal dengan satu titik knot, dua titik knot, tiga titik knot dan kombinasi knot.

4.3.1 Pemilihan Titik Knot dengan Satu Titik Knot

Estimasi model regresi nonparametrik spline dengan satu titik knot pada kasus jumlah kasus tuberkulosis di Jawa Timur adalah sebagai berikut

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - k_1)_+^1 + \hat{\beta}_3 x_2 + \hat{\beta}_4 (x_2 - k_2)_+^1 + \hat{\beta}_5 x_3 + \hat{\beta}_6 (x_3 - k_3)_+^1 + \hat{\beta}_7 x_4 + \hat{\beta}_8 (x_4 - k_4)_+^1 + \hat{\beta}_9 x_5 + \hat{\beta}_{10} (x_5 - k_5)_+^1 .$$

Tabel 4.2 merupakan sepuluh nilai GCV disekitar nilai GCV yang paling minimum untuk model regresi nonparametrik spline dengan satu knot.

Tabel 4.2 Nilai GCV Dengan Satu Titik Knot

X_1	X_2	X_3	X_4	X_5	GCV
82,92755	21,47143	3,263265	13,5638	58,74959	753902,4
84,1898	21,90429	3,346939	13,84182	59,74163	735337,7
85,45204	22,33714	3,430612	14,11984	60,73367	725544,8
86,71429	22,77	3,514286	14,39786	61,72571	700096,6
87,97653	23,20286	3,597959	14,67588	62,71776	661784,1
89,23878	23,63571	3,681633	14,9539	63,7098	638147,1
90,50102	24,06857	3,765306	15,23192	64,70184	650608,4
91,76327	24,50143	3,84898	15,50994	65,69388	658750
93,02551	24,93429	3,932653	15,78796	66,68592	687033,9
94,28776	25,36714	4,016327	16,06598	67,67796	692986,7

Berdasarkan Tabel 4.2 diketahui bahwa nilai GCV minimum untuk model regresi nonparametrik spline dengan satu knot adalah sebesar 638147,1. Nilai ini diperoleh dari satu titik knot optimal pada setiap variabel. Satu titik knot pada variabel persentase keluarga dengan kepemilikan sarana sanitasi dasar sehat (X_1) berada pada titik knot 89,23878; variabel prediktor persentase penduduk miskin (X_2) berada pada titik knot 23,63571; variabel persentase gizi buruk masyarakat (X_3) berada pada titik knot 3,681633; variabel persentase tenaga kesehatan terlatih TB (X_4) berada pada titik knot 14,9539 serta variabel persentase rumah tangga ber-PHBS (X_5) pada titik knot 63,7098.

4.3.2 Pemilihan Titik Knot dengan Dua Titik Knot

Setelah dilakukan pemilihan titik knot dengan satu titik knot, selanjutnya akan dilakukan pemilihan titik knot optimal menggunakan dua titik knot. Estimasi model regresi nonparametrik spline dengan dua titik knot pada jumlah kasus tuberkulosis di Jawa Timur adalah sebagai berikut

$$\begin{aligned} \hat{y} = & \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - k_1)_+^1 + \hat{\beta}_3 (x_1 - k_2)_+^1 + \hat{\beta}_4 x_2 + \\ & \hat{\beta}_5 (x_2 - k_3)_+^1 + \hat{\beta}_6 (x_2 - k_4)_+^1 + \hat{\beta}_7 x_3 + \hat{\beta}_8 (x_3 - k_5)_+^1 + \\ & \hat{\beta}_9 (x_3 - k_6)_+^1 + \hat{\beta}_{10} x_4 + \hat{\beta}_{11} (x_4 - k_7)_+^1 + \hat{\beta}_{12} (x_4 - k_8)_+^1 + \\ & \hat{\beta}_{13} x_5 + \hat{\beta}_{14} (x_5 - k_9)_+^1 + \hat{\beta}_{15} (x_5 - k_{10})_+^1. \end{aligned}$$

Tabel 4.3 menyajikan sepuluh nilai GCV di sekitar nilai GCV minimum untuk model regresi nonparametrik spline menggunakan dua titik knot.

Tabel 4.3 Nilai GCV Dengan Dua Titik Knot

X_1	X_2	X_3	X_4	X_5	GCV
90,50102	24,06857	3,765306	15,23192	64,70184	552026,8
91,76327	24,50143	3,84898	15,50994	65,69388	
90,50102	24,06857	3,765306	15,23192	64,70184	496007,5
93,02551	24,93429	3,932653	15,78796	66,68592	
90,50102	24,06857	3,765306	15,23192	64,70184	490169,8
94,28776	25,36714	4,016327	16,06598	67,67796	
90,50102	24,06857	3,765306	15,23192	64,70184	650608,4
95,55	25,8	4,1	16,344	68,67	
91,76327	24,50143	3,84898	15,50994	65,69388	490373
93,02551	24,93429	3,932653	15,78796	66,68592	
91,76327	24,50143	3,84898	15,50994	65,69388	489149
94,28776	25,36714	4,016327	16,06598	67,67796	
91,76327	24,50143	3,84898	15,50994	65,69388	658750
95,55	25,8	4,1	16,344	68,67	
93,02551	24,93429	3,932653	15,78796	66,68592	490041,5
94,28776	25,36714	4,016327	16,06598	67,67796	
93,02551	24,93429	3,932653	15,78796	66,68592	687033,9
95,55	25,8	4,1	16,344	68,67	
94,28776	25,36714	4,016327	16,06598	67,67796	692986,7
95,55	25,8	4,1	16,344	68,67	

Berdasarkan Tabel 4.3 diketahui bahwa nilai GCV terkecil yang diperoleh dari pemodelan regresi nonparametrik spline dengan dua titik knot adalah 489.149. Nilai GCV tersebut diperoleh pada titik knot 91,76327 dan 94,28776 untuk variabel X_1 , titik knot 24,50143 dan 25,36714 untuk X_2 , titik knot 3,84898 dan 4,016327 untuk X_3 , titik knot 15,50994 dan 16,06598 untuk X_4 , dan titik knot 65,69388 dan 67,67796 untuk X_5 . Apabila dibandingkan nilai GCV terkecil antara penggunaan satu titik knot dengan dua titik knot, terlihat bahwa nilai GCV dengan dua titik knot adalah lebih kecil. Maka regresi spline dengan dua titik knot lebih baik daripada menggunakan satu titik knot saja.

4.3.3 Pemilihan Titik Knot dengan Tiga Titik Knot

Estimasi model regresi nonparametrik spline dengan tiga titik knot untuk jumlah kasus tuberkulosis di Jawa Timur dengan variabel-variabel yang diduga mempengaruhi sebagai berikut

$$\begin{aligned}\hat{y} = & \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2(x_1 - k_1)_+^1 + \hat{\beta}_3(x_1 - k_2)_+^1 + \hat{\beta}_4(x_1 - k_3)_+^1 + \\ & \hat{\beta}_5 x_2 + \hat{\beta}_6(x_2 - k_4)_+^1 + \hat{\beta}_7(x_2 - k_5)_+^1 + \hat{\beta}_8(x_2 - k_6)_+^1 + \hat{\beta}_9 x_3 + \\ & \hat{\beta}_{10}(x_3 - k_7)_+^1 + \hat{\beta}_{11}(x_3 - k_8)_+^1 + \hat{\beta}_{12}(x_3 - k_9)_+^1 + \hat{\beta}_{13} x_4 + \\ & \hat{\beta}_{14}(x_4 - k_{10})_+^1 + \hat{\beta}_{15}(x_4 - k_{11})_+^1 + \hat{\beta}_{16}(x_4 - k_{12})_+^1 + \hat{\beta}_{17} x_5 + \\ & \hat{\beta}_{18}(x_5 - k_{13})_+^1 + \hat{\beta}_{19}(x_5 - k_{14})_+^1 + \hat{\beta}_{20}(x_5 - k_{15})_+^1.\end{aligned}$$

Berikut ini merupakan sepuluh nilai diantara nilai GCV minimum yang diperoleh dari regresi nonparametrik spline dengan tiga titik knot yang disajikan dalam Tabel 4.4 berikut ini.

Tabel 4.4 Nilai GCV Dengan Tiga Titik Knot

X_1	X_2	X_3	X_4	X_5	GCV
47,58469	9,351429	0,920408	5,779224	30,97245	400345,8
90,50102	24,06857	3,765306	15,23192	64,70184	
93,02551	24,93429	3,932653	15,78796	66,68592	
47,58469	9,351429	0,920408	5,779224	30,97245	385642,7
90,50102	24,06857	3,765306	15,23192	64,70184	
94,28776	25,36714	4,016327	16,06598	67,67796	
47,58469	9,351429	0,920408	5,779224	30,97245	373588,5
91,76327	24,50143	3,84898	15,50994	65,69388	
93,02551	24,93429	3,932653	15,78796	66,68592	
47,58469	9,351429	0,920408	5,779224	30,97245	372077,4
91,76327	24,50143	3,84898	15,50994	65,69388	
94,28776	25,36714	4,016327	16,06598	67,67796	
47,58469	9,351429	0,920408	5,779224	30,97245	371212,9
93,02551	24,93429	3,932653	15,78796	66,68592	
94,28776	25,36714	4,016327	16,06598	67,67796	
48,84694	9,784286	1,004082	6,057245	31,96449	1299061
50,10918	10,21714	1,087755	6,335265	32,95653	
51,37143	10,65	1,171429	6,613286	33,94857	

Tabel 4.4 Nilai GCV Dengan Tiga Titik Knot (Lanjutan)

X_1	X_2	X_3	X_4	X_5	GCV
48,84694	9,784286	1,004082	6,057245	31,96449	1226693
50,10918	10,21714	1,087755	6,335265	32,95653	
52,63367	11,08286	1,255102	6,891306	34,94061	
48,84694	9,784286	1,004082	6,057245	31,96449	1265443
50,10918	10,21714	1,087755	6,335265	32,95653	
53,89592	11,51571	1,338776	7,169327	35,93265	
48,84694	9,784286	1,004082	6,057245	31,96449	1297021
50,10918	10,21714	1,087755	6,335265	32,95653	
55,15816	11,94857	1,422449	7,447347	36,92469	
48,84694	9,784286	1,004082	6,057245	31,96449	1330897
50,10918	10,21714	1,087755	6,335265	32,95653	
56,42041	12,38143	1,506122	7,725367	37,91673	

Berdasarkan Tabel 4.4 diketahui bahwa nilai GCV terkecil yang diperoleh dari pemodelan regresi nonparametrik spline dengan tiga titik knot adalah 371212,9. Nilai GCV tersebut diperoleh pada titik knot 47,58469, 93,02551 dan 94,28776 untuk variabel X_1 , titik knot 9,351429, 24,93429 dan 25,36714 untuk variabel X_2 , titik knot 0,920408, 3,932653 dan 4,016327 untuk variabel X_3 , titik knot 5,779224, 15,78796 dan 16,06598 untuk variabel X_4 , dan titik knot 30,97245, 66,68592 dan 67,67796 untuk titik knot X_5 . Nilai GCV yang dihasilkan dengan menggunakan tiga titik knot lebih kecil daripada nilai GCV yang dihasilkan menggunakan satu titik knot atau dua titik knot.

4.3.4 Pemilihan Titik Knot dengan Kombinasi Knot

Pemilihan titik knot optimal dengan satu titik knot, dua titik knot serta tiga titik knot telah dilakukan. Selanjutnya, akan dilakukan pemilihan titik knot optimal dengan kombinasi titik knot karena terdapat kemungkinan bahwa setiap pola data memiliki jumlah titik knot optimal yang berbeda-beda.

Tabel 4.5 Nilai GCV Dengan Kombinasi Titik Knot

X_1	X_2	X_3	X_4	X_5	GCV
89,23878	9,351429	3,84898	15,50994	63,7098	684785,1
	24,93429	4,016327	16,06598		
	25,36714				
89,23878	9,351429	3,84898	15,50994	65,69388	475360,7
	24,93429	4,016327	16,06598	67,67796	
	25,36714				
89,23878	9,351429	3,84898	5,779224	65,69388	276418,0578
	24,93429	4,016327	15,78796	67,67796	
	25,36714		16,06598		
89,23878	9,351429	3,84898	15,50994	30,97245	504080,1
	24,93429	4,016327	16,06598	66,68592	
	25,36714			67,67796	
89,23878	9,351429	3,84898	5,779224	63,7098	595520,6
	24,93429	4,016327	15,78796		
	25,36714		16,06598		

Berdasarkan Tabel 4.5 diketahui bahwa nilai GCV minimum dari kombinasi knot adalah 276418,0578. Nilai GCV tersebut dihasilkan apabila digunakan kombinasi knot (1,3,2,3,2) pada masing-masing variabel X_1 , X_2 , X_3 , X_4 dan X_5 . Titik knot yang digunakan adalah 89,23878 untuk X_1 , 9,351429, 24,93429 dan 25,36714 untuk X_2 , 3,84898 dan 4,016327 untuk X_3 , 5,779224, 15,78796 dan 16,06598 untuk X_4 dan 65,69388 dan 67,67796 untuk X_5 . Selanjutnya akan dilakukan pemilihan model terbaik berdasarkan nilai GCV yang paling minimum.

4.3.5 Pemilihan Model Terbaik

Setelah didapatkan nilai GCV minimum pada pemilihan titik knot optimal dengan satu titik knot, dua titik knot, tiga titik knot dan kombinasi knot, selanjutnya dilakukan pemilihan model terbaik. Pemilihan model terbaik dilakukan dengan membandingkan nilai dari GCV yang paling minimum. Tabel 4.6 merupakan tabel yang menunjukkan nilai GCV terkecil dari masing-masing model.

Tabel 4.6 Nilai GCV Minimum Setiap Pemilihan Titik Knot Optimal

Banyak Titik Knot Optimal	Nilai GCV Minimum
Satu Titik Knot	638147,0798
Dua Titik Knot	489149,0305
Tiga Titik Knot	371212,9481
Kombinasi Knot (1,3,2,3,2)	276418,0578 (*)

Berdasarkan Tabel 4.6 diketahui bahwa pemodelan yang menghasilkan nilai GCV paling minimum merupakan pemodelan regresi nonparametrik spline dengan menggunakan kombinasi knot (1,3,2,3,2). Oleh karena itu, diputuskan bahwa model terbaik yang akan dipilih adalah model regresi nonparametrik spline dengan menggunakan kombinasi knot 1,3,2,3,2.

4.3.6 Penaksiran Parameter Model Regresi Nonparametrik Spline

Berdasarkan kriteria pemilihan model terbaik diketahui bahwa nilai GCV paling minimum dihasilkan oleh model regresi nonparametrik spline dengan kombinasi titik knot 1,3,2,3,2. Estimasi parameter didapatkan menggunakan metode *Ordinary Least Square* (OLS). Berikut ini merupakan model regresi nonparametrik spline terbaik yang dihasilkan.

$$\begin{aligned} \hat{y} = & 2184,86 - 2,564X_1 - 75,863(X_1 - 89,238)_+^1 + 141,526X_2 + \\ & - 128,430(X_2 - 9,351)_+^1 - 21,847(X_2 - 24,934)_+^1 + \\ & - 10,923(X_2 - 25,367)_+^1 + 176,203X_3 - 6129,817(X_3 - 3,848)_+^1 + \\ & - 2043,264(X_3 - 4,0163) - 595,473X_4 + \\ & + 649,170(X_4 - 5,7792)_+^1 - 368,3163(X_4 - 15,7879)_+^1 + \\ & - 184,158(X_4 - 16,0659) + 12,7073X_5 + \\ & + 2647,442(X_5 - 65,6938)_+^1 - 7652,740(X_5 - 67,6779)_+^1. \end{aligned}$$

4.4 Pengujian Signifikansi Parameter Model Regresi Nonparametrik Spline

Setelah mendapatkan model terbaik, maka dilakukan pengujian signifikansi parameter model yang bertujuan untuk mengetahui apakah variabel-variabel prediktor berpengaruh secara signifikan terhadap jumlah kasus tuberkulosis di Jawa Timur atau tidak. Pengujian signifikansi parameter dilakukan

secara serentak terlebih dahulu. Apabila hasil uji signifikansi parameter secara serentak menunjukkan bahwa terdapat minimal satu parameter yang signifikan, maka pengujian dapat dilanjutkan secara individu/parsial.

4.4.1 Uji Serentak

Tujuan dilakukan uji serentak adalah menguji signifikansi dari seluruh parameter yang terdapat dalam model secara keseluruhan (simultan). Berikut ini hasil analisis ragam model regresi nonparametrik spline yang disajikan dalam Tabel 4.7.

Tabel 4.7 Analisis Ragam Uji Serentak

Sumber Variasi	Derajat Bebas (df)	Jumlah Kuadrat (SS)	Rataan Kuadrat (MS)	F_{hitung}	$p-value$
Regresi	16	21078174	1317386		
Error	21	4189916	199519,8	6,602783	$5,09448 \times 10^{-5}$
Total	37	25268090			

Berdasarkan Tabel 4.7 diketahui bahwa nilai $p-value$ sebesar $5,09448 \times 10^{-5}$. Nilai ini adalah lebih kecil daripada α (0,05). Disamping itu, nilai $F_{hitung} > F_{(0,05;16;21)}$ yaitu $6,602783 > 2,1562$ yang menunjukkan bahwa H_0 ditolak, artinya minimal terdapat satu parameter yang signifikan dalam model. Untuk mengetahui parameter manakah yang signifikan terhadap model, dilakukan uji individu.

4.4.2 Uji Individu

Pengujian parameter secara individu dilakukan untuk mengetahui variabel yang berpengaruh secara signifikan terhadap jumlah kasus tuberkulosis di Jawa Timur. Berikut ini merupakan Tabel 4.8 yang menunjukkan hasil pengujian individu pada masing-masing parameter.

Tabel 4.8 Hasil Uji Individu

Variabel	Parameter	Estimasi Parameter	<i>p-value</i>	t-hitung	Keputusan
Constant	β_0	2184,864056	0,05915988	1,99518718	Tidak signifikan
X_1	β_1	-2,564361	0,7793626	-0,28377388	Tidak signifikan
	β_2	-75,863216	0,2688836	-1,13568625	Tidak signifikan
X_2	β_3	141,526198	0,05166073	2,06333110	Tidak signifikan
	β_4	-128,430085	0,1209485	-1,61632538	Tidak signifikan
	β_5	-21,847054	0,9664244	-0,04259805	Tidak signifikan
	β_6	-10,923653	0,9664244	-0,04259805	Tidak signifikan
X_3	β_7	176,203967	0,1426633	1,52305267	Signifikan
	β_8	-6129,817085	0,005388391	-3,10278044	Signifikan
	β_9	-2043,264222	0,005388391	-3,10278044	Signifikan
X_4	β_{10}	-595,473536	$9,0206 \times 10^{-5}$	-4,82729856	Signifikan
	β_{11}	649,170892	0,0003451	4,26453351	Signifikan
	β_{12}	-368,316391	0,6015729	-0,53014464	Tidak Signifikan
	β_{13}	-184,158195	0,6015729	-0,53014464	Tidak Signifikan
X_5	β_{14}	12,707361	0,06177257	1,97324937	Tidak Signifikan
	β_{15}	2647,442589	$1,0009 \times 10^{-5}$	5,76871612	Signifikan
	β_{16}	-7652,740754	$1,4985 \times 10^{-5}$	-5,59322919	Signifikan

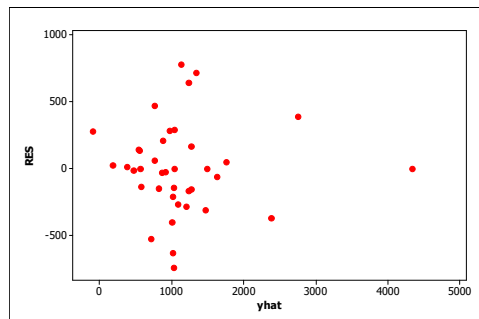
Apabila nilai *p-value* kurang dari α yaitu 0,05 maka estimasi parameter signifikan. Berdasarkan Tabel 4.9 diketahui bahwa dengan tingkat kepercayaan 95%, terdapat tiga variabel yang berpengaruh secara signifikan terhadap model yaitu persentase gizi buruk masyarakat (X_3), persentase tenaga kesehatan terlatih TB (X_4), dan persentase rumah tangga ber-PHBS (X_5). Namun, terdapat dua variabel yang tidak berpengaruh secara signifikan terhadap model yaitu variabel X_1 yaitu persentase keluarga dengan kepemilikan sarana sanitasi dasar sehat dan persentase penduduk miskin (X_2) tidak berpengaruh secara signifikan dalam model.

4.5 Pengujian Asumsi Residual

Penggunaan metode regresi nonparametrik spline harus memenuhi asumsi residual IIDN. Hal ini berarti bahwa residual dari model regresi nonparametrik yang telah diperoleh harus memenuhi asumsi identik, independen dan berdistribusi normal. Oleh karena itu, akan dilakukan tiga pengujian terhadap residual model regresi nonparametrik spline sebagai berikut.

4.5.1 Pemeriksaan Asumsi Identik

Residual dari hasil regresi nonparametrik spline harus memenuhi asumsi identik yang artinya tidak terjadi kasus heterokedastisitas pada residual. Dengan kata lain, variansi residual dari model harus homogen. Pemeriksaan asumsi residual identik dapat dilakukan dengan dua cara yaitu berdasarkan visual menggunakan *scatterplot* atau secara matematis menggunakan uji Glejser. Gambar 4.7 dibawah merupakan *scatterplot* antara residual dengan \hat{y} untuk mendeteksi secara visual terjadinya heteroskedastisitas pada residual.



Gambar 4.7 *Scatterplot* antara Residual dengan \hat{y}

Berdasarkan Gambar 4.7 diketahui bahwa hasil *scatterplot* antara residual dengan \hat{y} tidak membentuk pola tertentu atau berpola acak. Hal ini menunjukkan bahwa tidak terjadi kasus heteroskedastisitas. Selain mendeteksi secara visual, akan dilakukan uji Glejser untuk mendeteksi kasus heteroskedastisitas. Berikut ini merupakan Tabel 4.9 Hasil analisis ragam uji Glejser.

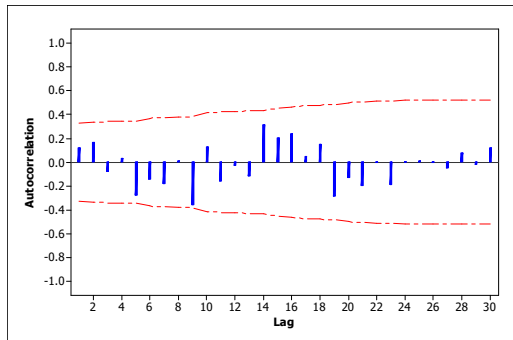
Tabel 4.9 Analisis Ragam Uji Glejser

Sumber Variasi	Derajat Bebas (df)	Jumlah Kuadrat (SS)	Rataan Kuadrat (MS)	F_{hitung}	$p-value$
Regresi	16	827030,4	51689,4		
Error	21	1104173	52579,64	0,9830687	0,5058166
Total	37	1931203			

Berdasarkan Tabel 4.9 diketahui bahwa nilai $p-value$ adalah sebesar 0,5058166 yang lebih besar daripada α yaitu 0,05 serta nilai $F_{hitung} < F_{(0,05;16,21)}$ yaitu $0,9830687 < 2,1562$ sehingga gagal tolak H_0 , artinya variansi residual homogen atau tidak terjadi heterkodastisitas dalam model. Baik pengujian secara visual menggunakan *scatterplot* ataupun berdasarkan uji Glejser, diperoleh hasil yang sama yaitu residual model regresi nonparametrik spline memenuhi asumsi identik.

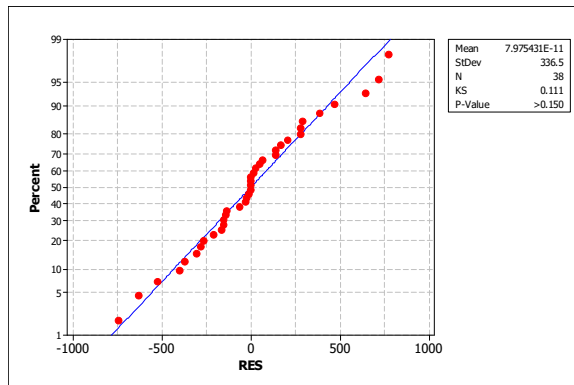
4.5.2 Pemeriksaan Asumsi Independen

Asumsi berikutnya yang harus terpenuhi adalah residual yang independen. Residual independen adalah tidak terjadi autokorelasi antar residual. Plot *Autocorrelation Function* (ACF) dapat digunakan untuk mengetahui ada atau tidaknya autokorelasi antar residual. Gambar 4.8 menunjukkan bahwa tidak terlihat adanya autokorelasi yang keluar batas interval konfidensi. Hal ini mengindikasikan bahwa asumsi independen pada residual model regresi nonparametrik spline telah terpenuhi.

**Gambar 4.8** Plot ACF Residual

4.5.3 Pemeriksaan Asumsi Distribusi Normal

Asumsi terakhir yang harus dipenuhi adalah residual yang dihasilkan berdistribusi normal. Pengujian apakah residual berdistribusi normal atau tidak menggunakan uji *Kolmogorov-Smirnov*. Berdasarkan Gambar 4.9 diketahui bahwa nilai *Kolmogorov-Smirnov* sebesar 0,111 dimana nilai ini lebih kecil dibandingkan $q_{(1-\alpha)}$ yaitu 0,231 dan *p-value* adalah $> 0,150$ sehingga gagal tolak H_0 . Hal ini menunjukkan bahwa residual model regresi nonparametrik spline telah memenuhi asumsi distribusi normal.



Gambar 4.9 Hasil Uji *Kolmogorov-Smirnov*

4.6 Nilai Koefisien Determinasi (R^2)

Berikut ini merupakan penghitungan R^2 dengan $SS_{regresi}$ dan SS_{total} diperoleh berdasarkan Tabel 4.10.

$$\begin{aligned}
 R^2 &= \frac{SS_{regresi}}{SS_{total}} \times 100\% \\
 &= 83,41815309 \% \\
 &\approx 83,42 \%
 \end{aligned}$$

Berdasarkan perhitungan tersebut didapatkan R^2 sebesar 83,42%. Hal ini menunjukkan bahwa model tersebut mampu menjelaskan keragaman jumlah kasus tuberkulosis di Jawa Timur sebesar 83,42%, sedangkan sisanya dijelaskan oleh variabel lain.

4.7 Interpretasi Model Regresi Nonparamterik Spline

Setelah dilakukan pengujian asumsi pada residual model regresi nonparametrik spline dan diperoleh bahwa semua asumsi tersebut terpenuhi, selanjutnya akan dilakukan intepretasi model. Model regresi nonparametrik spline terbaik yang telah memenuhi asumsi IIDN $(0, \sigma^2)$ untuk data jumlah kasus tuberkulosis di Jawa Timur adalah sebagai berikut

$$\begin{aligned}\hat{y} = & 2184,86 - 2,564X_1 - 75,863(X_1 - 89,238)_+^1 + 141,526X_2 + \\ & - 128,430(X_2 - 9,351)_+^1 - 21,847(X_2 - 24,934)_+^1 + \\ & - 10,923(X_2 - 25,367)_+^1 + 176,203X_3 - 6129,817(X_3 - 3,848)_+^1 + \\ & - 2043,264(X_3 - 4,0163)_+^1 - 595,473X_4 + \\ & + 649,170(X_4 - 5,7792)_+^1 - 368,3163(X_4 - 15,7879)_+^1 + \\ & - 184,158(X_4 - 16,0659)_+^1 + 12,7073X_5 + \\ & + 2647,442(X_5 - 65,6938)_+^1 - 7652,740(X_5 - 67,6779)_+^1.\end{aligned}$$

Dari model tersebut dapat diinterpretasikan sebagai berikut:

1. Apabila variabel X_2 , X_3 , X_4 dan X_5 dianggap konstan, maka pengaruh persentase keluarga dengan kepemilikan sarana sanitasi dasar sehat (X_1) terhadap jumlah kasus tuberkulosis di Jawa Timur adalah

$$\begin{aligned}\hat{y} = & 2184,86 - 2,564X_1 - 75,863(X_1 - 89,238)_+^1 \\ = & \begin{cases} 2184,86 - 2,564X_1, & X_1 < 89,238 \\ 8954,722 - 754,427 X_1, & X_1 \geq 89,238 \end{cases}\end{aligned}$$

Pada model tersebut dapat diinterpretasikan dengan visualisasi peta sebagai berikut.



Gambar 4.10 Peta Persebaran Jumlah Kasus TB Berdasarkan Persentase Keluarga dengan Kepemilikan Sarana Sanitasi Dasar Sehat

Apabila persentase keluarga dengan kepemilikan sarana sanitasi dasar sehat kurang dari 89,238% yaitu pada Kabupaten/Kota selain Kabupaten Pacitan, Kabupaten Gresik, Kabupaten Ngawi, Kota Kediri, Kota Surabaya, dan Kabupaten Magetan, maka jika persentase keluarga dengan kepemilikan sarana sanitasi dasar sehat naik sebesar 1 persen, maka kasus tuberkulosis akan turun sebesar 2,564 kasus. Dan apabila persentase keluarga dengan kepemilikan sarana sanitasi dasar sehat lebih dari 89,238% yaitu pada Kabupaten Pacitan, Kabupaten Gresik, Kabupaten Ngawi, Kota Kediri, Kota Surabaya, dan Kabupaten Magetan, maka apabila persentase keluarga dengan kepemilikan sarana sanitasi dasar sehat naik sebesar 1 persen, maka kasus tuberkulosis akan turun sebesar 754,427.

2. Apabila variabel X_1 , X_3 , X_4 dan X_5 dianggap konstan, maka pengaruh persentase penduduk miskin (X_2) terhadap jumlah kasus tuberkulosis di Jawa Timur adalah

$$\hat{y} = 2184,86 + 141,526X_2 - 128,430(X_2 - 9,351)^{\frac{1}{4}} - 21,847(X_2 - 24,934)^{\frac{1}{4}} - 10,923(X_2 - 25,367)^{\frac{1}{4}}$$

$$= \begin{cases} 2184,86 + 141,526X_2, & X_2 < 9,351 \\ 3385,8089 + 13,096 X_2, & 9,351 \leq X_2 < 24,934 \\ 3930,5410 - 8,751 X_2, & 24,934 \leq X_2 < 25,367 \\ 4207,6247 - 19,674 X_2, & X_2 \geq 25,367 \end{cases}$$

Pada model tersebut dapat diinterpretasikan dengan visualisasi peta sebagai berikut.



Gambar 4.11 Peta Persebaran Jumlah Kasus TB Berdasarkan Persentase Penduduk Miskin

Apabila persentase penduduk miskin kurang dari 9,351%, yaitu pada Kota Batu, Kota Malang, Kota Madiun, Kota Surabaya, Kabupaten Sidoarjo, Kota Mojokerto, Kota Blitar, Kota Pasuruan, Kota Kediri, Kota Probolinggo, Kabupaten Tulungagung, dan Kabupaten Banyuwangi, maka apabila persentase penduduk miskin naik sebesar 1 persen, maka kasus tuberculosis akan naik sebesar 141,526 kasus. Apabila persentase penduduk miskin berkisar antara 9,351 dan 24,934% yaitu pada sebanyak 25 kabupaten/kota di Jawa Timur, maka apabila persentase penduduk miskin naik sebesar 1 persen, maka kasus tuberculosis akan naik sebesar 13,096. Apabila persentase penduduk miskin adalah lebih besar dari 25,367% yaitu pada Kabupaten Sampang, maka apabila persentase penduduk miskin bertambah 1 persen, maka kasus tuberculosis akan turun sebesar 19,674.

3. Apabila variabel X_1 , X_2 , X_4 dan X_5 dianggap konstan, maka pengaruh persentase gizi buruk masyarakat (X_3) terhadap jumlah kasus tuberculosis di Jawa Timur adalah.

$$\begin{aligned} \hat{y} &= 2184,86 + 176,203X_3 - 6129,817(X_3 - 3,848)^{\frac{1}{4}} \\ &\quad - 2043,264(X_3 - 4,0163)^{\frac{1}{4}} \\ &= \begin{cases} 2184,86 + 176,203 X_3, & X_3 < 3,848 \\ 25772,395 - 5953,614 X_3, & 3,848 \leq X_3 < 4,0163 \\ 33978,7562 - 7996,878 X_3, & X_3 \geq 4,0163 \end{cases} \end{aligned}$$

Pada model tersebut dapat diinterpretasikan dengan visualisasi peta sebagai berikut.



Gambar 4.12 Peta Persebaran Jumlah Kasus TB Berdasarkan Persentase Gizi Buruk Masyarakat

Berdasarkan model tersebut, dapat diinterpretasikan bahwa jika persentase tenaga kesehatan terlatih TB dalam suatu provinsi

kurang dari 5,7792% yaitu untuk Kabupaten Sidoarjo, Kota Pasuruan, Kabupaten Jember, Kota Malang, Kabupaten Malang dan Kabupaten Jombang, maka apabila persentase tenaga kesehatan terlatih TB naik sebesar 1 persen, maka kasus tuberkulosis akan naik sebesar 595,473. Kemudian, apabila persentase tenaga kesehatan terlatih TB berkisar antara 5,7792 dan 15,7879%, maka apabila persentase tenaga kesehatan terlatih TB naik sebesar 1 persen, maka akan menaikkan kasus tuberkulosis sebesar 1244,643 yaitu sebanyak 30 kabupen/kota di Jawa Timur, apabila persentase tenaga kesehatan terlatih TB adalah lebih dari 16,0659% yaitu di Kabupaten Tulungagung dan Kabupaten Sumenep, maka apabila persentase tenaga kesehatan terlatih TB naik sebesar 1 persen, akan meningkatkan kasus tuberkulosis sebanyak 692,1687.

5. Apabila variabel X_1 , X_2 , X_3 dan X_4 dianggap konstan, maka pengaruh persentase rumah tangga ber-PHBS (X_5) terhadap jumlah kasus tuberkulosis di Jawa Timur adalah

$$\begin{aligned} \hat{y} &= 2184,86 + 12,7073X_5 + 2647,442(X_5 - 65,6938)^{\frac{1}{4}} \\ &\quad - 7652,740(X_5 - 67,6779)^{\frac{1}{4}} \\ &= \begin{cases} 2184,86 + 12,7073X_5, & X_5 < 65,6938 \\ -171735,66 + 2660,1493 X_5, & 65,6938 \leq X_5 < 67,6779 \\ 346185,712 - 4992,590 X_5, & X_5 \geq 67,6779 \end{cases} \end{aligned}$$

Pada model tersebut dapat diinterpretasikan dengan visualisasi peta sebagai berikut.



Gambar 4.14 Peta Persebaran Jumlah Kasus TB Berdasarkan Persentase Rumah Tangga Ber-PHBS

Berdasarkan model tersebut, dapat diinterpretasikan bahwa jika persentase rumah tangga ber-PHBS kurang dari 65,6938% yaitu pada semua kabupaten/kota di Jawa Timur selain Kota Surabaya dan Kabupaten Gresik, maka apabila persentase rumah tangga ber-PHBS naik sebesar 1 persen, maka kasus tuberkulosis akan naik sebesar 12,7073. Selanjutnya apabila persentase rumah tangga ber-PHBS lebih dari 67,6779% yaitu pada Kota Surabaya dan Kabupaten Gresik, maka apabila persentase rumah tangga ber-PHBS naik sebesar 1 persen, maka kasus tuberkulosis akan turun sebesar 4992,590.

(halaman ini sengaja dikosongkan)

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, dapat disimpulkan beberapa hal sebagai berikut.

1. Jumlah kasus tuberkulosis di Jawa Timur pada tahun 2014 sebanyak 42.458 kasus dengan kasus BTA positif baru sebanyak 22.866 kasus. Kota Surabaya merupakan kota yang memiliki jumlah kasus tuberkulosis terbanyak yaitu sebesar 4.338 kasus. Tingginya jumlah kasus tuberkulosis di dapat disebabkan karena Kota Surabaya merupakan salah satu kota dengan tingkat kepadatan penduduk tertinggi di Jawa Timur yaitu sebesar 8683 jiwa/km². Kabupaten/kota dengan jumlah kasus tuberkulosis terbanyak kedua adalah Kabupaten Jember yaitu sebanyak 3.139 kasus. Kota dengan jumlah kasus tuberkulosis terendah adalah Kota Batu yaitu sebanyak 176 kasus.
2. Model regresi nonparametrik spline terbaik untuk pemodelan jumlah kasus tuberkulosis di Jawa Timur adalah dengan menggunakan kombinasi knot 1,3,2,3,2. Model ini mempunyai nilai koefisien determinasi (R^2) sebesar 83,42% dengan tiga variabel yang berpengaruh secara signifikan yaitu persentase gizi buruk masyarakat, persentase tenaga kesehatan terlatih TB, dan persentase rumah tangga ber-PHBS. Model regresi nonparametrik spline yang diperoleh adalah sebagai berikut.

$$\begin{aligned}\hat{y} = & 2184,86 - 2,564X_1 - 75,863(X_1 - 89,238)_+^1 + 141,526X_2 + \\ & - 128,430(X_2 - 9,351)_+^1 - 21,847(X_2 - 24,934)_+^1 + \\ & - 10,923(X_2 - 25,367)_+^1 + 176,203X_3 - 6129,817(X_3 - 3,848)_+^1 \\ & - 2043,264(X_3 - 4,0163)_+^1 - 595,473X_4 + \\ & + 649,170(X_4 - 5,7792)_+^1 - 368,3163(X_4 - 15,7879)_+^1 + \\ & - 184,158(X_4 - 16,0659)_+^1 + 12,7073X_5 + \\ & + 2647,442(X_5 - 65,6938)_+^1 - 7652,740(X_5 - 67,6779)_+^1.\end{aligned}$$

5.2 Saran

Berdasarkan hasil penelitian mengenai jumlah kasus tuberkulosis di Jawa Timur dengan menggunakan regresi nonparametrik spline didapatkan bahwa persentase variabel gizi buruk masyarakat, persentase variabel tenaga kesehatan terlatih TB, dan persentase rumah tangga ber-PHBS berpengaruh signifikan maka disarankan kepada khususnya dinas kesehatan provinsi Jawa Timur untuk mensosialisasikan mengenai pentingnya kebersihan lingkungan bagi kesehatan. Sedangkan untuk penelitian selanjutnya diharapkan perlu adanya pengembangan menggunakan empat titik knot karena pada penelitian ini masih terbatas menggunakan regresi nonparametrik spline linier hanya sampai tiga knot.

DAFTAR PUSTAKA

- Amin, M., Alsagaff, H. & Saleh, W., 1989. *Pengantar Ilmu Penyakit Paru*. Surabaya: Airlangga University Press.
- Budiantara, I. N., 2006. Model Spline dengan Knots Optimal. *Jurnal Ilmu Dasar*, Volume Vol 7, pp. Hal 77-85.
- Budiantara, I. N., 2009. *Spline dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Mendatang*. Surabaya: ITS Press.
- Daniel, W. W., 1990. *Applied Nonparametric Statistical Method (2nd ed)*. Boston: PWS-Kent Publishing Company.
- Departemen Kesehatan Republik Indonesia, 2001. *Pedoman Nasional dan Penanggulangan TB*. Jakarta: Depkes RI.
- Departemen Kesehatan Republik Indonesia, 2009. *Keputusan Menteri Kesehatan RI tentang Pedoman Penanggulangan TB*. Jakarta: Depkes RI.
- Dinas Kesehatan Provinsi Jawa Timur, 2013. *Profil Kesehatan Provinsi Jawa Timur Tahun 2012*. Surabaya: Dinas Kesehatan Provinsi Jawa Timur.
- Draper, N. & Smith, H., 1992. *Analisis regresi Terapan Edisi Kedua Diterjemahkan oleh Bambang Sumantri*. Jakarta: PT. Gramedia Pustaka Utama.
- Dwikentarti, F., 2010. *Analisis Faktor-Faktor yang Mempengaruhi Penyakit Tuberkulosis pada Pasien dengan Regresi Logistik Multinomial*. Semarang: Tugas Akhir Jurusan Statistika Universitas Diponegoro.
- Eubank, R., 1988. *Spline Smoothing and Nonparametric Regression*. New York: Marcel Dekker Inc..
- Fatimah, S., 2008. *Faktor Kesehatan Lingkungan Rumah yang Berhubungan dengan Kejadian TB Paru di Kabupaten Cilacap*. Semarang: Thesis Jurusan Magister Kesehatan Lingkungan Universitas Diponegoro.
- Hasyim, M., 2010. *Pemodelan Angka Kejadian Penyakit Infeksi Tuberkulosis Paru (TB Paru) di Kabupaten Sorong Selatan (Provinsi Papua Barat) dengan Pendekatan Multivariate*

- Adaptive Regression Spline (MARS)*. Surabaya: Tugas Akhir Jurusan Statistika FMIPA-ITS Surabaya.
- Kementrian Kesehatan RI, 2011. *Strategi Nasional Pengendalian Tuberkulosis di Indonesia 2010-2014*. Jakarta: Kementrian Kesehatan RI Direktorat Jendral Pengendalian Penyakit dan Penyehatan Lingkungan.
- Lestari, R. D., 2014. *Pemodelan Faktor-Faktor yang Mempengaruhi Jumlah Kasus Penyakit Tuberkulosis di Jawa Timur dengan Pendekatan Generalized Poisson Regression (GPR) dan Geographically Weighted Poisson Regression (GWPR)*. Surabaya: Tugas Akhir Jurusan Statistika ITS Surabaya.
- McClave, J., Bendon, P. & Sincich, T., 2010. *Statistics for Business and Economics Eleventh Edition*. Jakarta: Diterjemahkan oleh Penerbit Erlangga.
- Puspita, E., 2014. *Analisis Regresi Logistik Biner Pada Faktor Resiko yang Mempengaruhi Penderita Penyakit TB Paru di RSUD Haji Surabaya*. Surabaya: Tugas Akhir Jurusan Statistika ITS Surabaya.
- Siagian, V., 2011. *Program Nasional Penanggulangan Tuberkulosis*. Manado: Kongres Ilmiah dan Rakernas IAI.
- Sukmawati, E., 2010. *Analisis Faktor-Faktor yang Mempengaruhi Laju Kesembuhan Penderita Penyakit Tuberkulosis di RSUD Ibnu Sina Kabupaten Gresik*. Surabaya: Tugas Akhir Jurusan Statistika ITS Surabaya.
- Walpole, R. E., 1995. *Pengantar Statistika*. Diterjemahkan oleh Bambang Sumantri. Jakarta: PT. Gramedia Pusaka Utama.
- Wei, W. W., 2006. *Time Series Analysis Univariate and Multivariate Methods*. United States: Pearson Education, Inc.

LAMPIRAN

Lampiran 1. Data Jumlah Kasus Tuberkulosis di Jawa Timur dengan Faktor-Faktor Yang Mempengaruhi Tahun 2014

PROVINSI	Y	X ₁	X ₂	X ₃	X ₄	X ₅
Pacitan	181	91,75	16,18	0,63	7,657	60,26
Ponorogo	813	72,09	11,53	1,48	7,642	63,54
Trenggalek	382	63,22	13,1	1,66	13,527	27,73
Tulungagung	819	80,79	8,75	1,06	16,344	38,12
Blitar	658	74,54	10,22	0,91	9,1	46,00
Kediri	1240	71,79	12,77	1,39	7,223	55,83
Malang	1801	80,87	11,07	1,21	3,604	28,33
Lumajang	1077	74,79	11,75	2,30	6,101	42,51
Jember	3139	53,87	11,28	3,21	3,393	63,98
Banyuwangi	1869	67,90	9,29	2,88	11,297	42,40
Bondowoso	804	41,74	14,76	3,34	8,389	20,06
Situbondo	883	33,70	13,15	2,59	10,575	22,07
Probolinggo	1329	49,00	20,44	2,56	9,68	22,99
Pasuruan	2053	56,75	10,86	2,89	12,189	42,67
Sidoarjo	2009	67,91	6,4	1,69	2,721	61,98
Mojokerto	832	70,53	10,56	1,92	7,411	42,23
Jombang	1160	74,10	10,8	2,32	4,966	53,38
Nganjuk	595	78,51	13,14	2,63	8,829	35,78
Madiun	911	78,09	12,04	1,37	9,979	64,98
Magetan	682	95,55	11,8	1,37	8,333	62,29
Ngawi	676	92,98	14,88	2,37	6,645	35,80
Bojonegoro	1223	68,31	15,48	0,00	6,789	57,41
Tuban	879	63,64	16,64	1,76	6,466	44,00
Lamongan	1905	81,78	15,68	1,82	7,513	61,04
Gresik	1489	92,12	13,41	2,15	11,086	68,67
Bangkalan	1068	68,03	22,38	3,04	7,657	42,00
Sampang	1035	61,34	25,8	2,58	7,642	29,82
Pamekasan	1110	59,72	17,74	3,03	13,527	22,88
Sumenep	1562	47,09	20,49	2,96	16,344	55,00

PROVINSI	Y	X ₁	X ₂	X ₃	X ₄	X ₅
K. Kediri	450	94,15	7,95	2,05	9,1	52,61
K. Blitar	206	84,76	7,15	0,30	7,223	40,39
K. Malang	1435	84,09	4,8	1,37	3,604	41,36
K. Probolinggo	282	70,05	8,37	2,88	6,101	59,22
⋮	⋮	⋮	⋮	⋮	⋮	⋮
K. Batu	176	87,50	4,59	0,93	9,68	28,14

Keterangan:

Y = Jumlah Kasus Tuberkulosisc

X_1 = Persentase Keluarga Dengan Kepemilikan Sarana Sanitasi Dasar Sehat

X_2 = Persentase Penduduk Miskin

X_3 = Persentase Gizi Buruk Masyarakat

X_4 = Persentase Tenaga Kesehatan Terlatih TB

X_5 = Persentase Rumah Tangga Ber-PHBS

Lampiran 2. Program Pemilihan Titik Knot Optimal dengan Satu Titik Knot Menggunakan *Software R*

```
GCV1=function(para)
{
  data=read.table("d://datatb.txt")
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-para-1
  dataA=data[, (para+2):q]
  F=matrix(0,nrow=p,ncol=p)
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot1=matrix(ncol=m,nrow=nk)
  for (i in (1:m))
  {
    for (j in (1:nk))
    {
      a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
      knot1[j,i]=a[j]
    }
  }
  a1=length(knot1[,1])
  knot1=knot1[2:(a1-1),]
  aa=rep(1,p)
  data1=matrix(ncol=m,nrow=p)
  data2=data[,2:q]
  a2=nrow(knot1)
  GCV=rep(NA,a2)
  Rsq=rep(NA,a2)
  for (i in 1:a2)
  {
    for (j in 1:m)
    {
      for (k in 1:p)
```

```

    {
if (data[k,(j+para+1)]<knot1[i,j]) data1[k,j]=0 else
    data1[k,j]=data[k,(j+para+1)]-knot1[i,j]
    }
    }
mx=cbind(aa,data2,data1)
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in (1:p))
{
    sum=(data[r,1]-yhat[r,])^2
    sum1=(yhat[r,]-mean(data[,1]))^2
    SSE=SSE+sum
    SSR=SSR+sum1
}
Rsqr[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsqr=as.matrix(Rsqr)
cat("=====", "\n")
cat("Nilai Knot dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print (knot1)
cat("=====", "\n")
cat("Rsqr dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print (Rsqr)
cat("=====", "\n")

```

```

cat("HASIL GCV dengan Spline linear 1 knot","\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
print(max(Rsq))
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 1 knot","\n")
cat("=====", "\n")
cat(" GCV =",s1,"\n")
write.table(GCV,file="d:/output GCV1.csv",sep=";")
write.table(Rsq,file="d:/output Rsq1.csv",sep=";")
write.table(knot1,file="d:/output knot1.csv",sep=";")
}

```

Lampiran 3. Program Pemilihan Titik Knot Optimal dengan Dua Titik Knot Menggunakan *Software R*

```
GCV2=function()
{
  data=read.table("D:/datatb.txt", header=FALSE)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-1
  F=matrix(0,nrow=p,ncol=p)
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot=matrix(ncol=m,nrow=nk)
  for (i in (1:m))
  {
    for (j in (1:nk))
    {
      a=seq(min(data[,i+1]),max(data[,i+1]),length.out=50)
      knot[j,i]=a[j]
    }
  }
  z=(nk*(nk-1)/2)
  knot2=cbind(rep(NA,(z+1)))
  for (i in (1:m))
  {
    knot1=rbind(rep(NA,2))
    for (j in 1:(nk-1))
    {
      for (k in (j+1):nk)
      {
        xx=cbind(knot[j,i],knot[k,i])
        knot1=rbind(knot1,xx)
      }
    }
  }
  knot2=cbind(knot2,knot1)
}
```

```

knot2=knot2[2:(z+1),2:(2*m+1)]
aa=rep(1,p)
data2=matrix(ncol=(2*m),nrow=p)
data1=data[,2:q]
a1=length(knot2[,1])
GCV=rep(NA,a1)
Rsqr=rep(NA,a1)
for (i in 1:a1)
{
  for (j in 1:(2*m))
  {
    if (mod(j,2)==1) b=floor(j/2)+1 else b=j/2
    for (k in 1:p)
    {
      if (data1[k,b]<knot2[i,j]) data2[k,j]=0 else
      data2[k,j]=data1[k,b]-knot2[i,j]
    }
  }
  mx=cbind(aa,data1,data2)
  mx=as.matrix(mx)
  C=pinv(t(mx)%*%mx)
  B=C%*%(t(mx)%*%data[,1])
  yhat=mx%*%B
  SSE=0
  SSR=0
  for (r in (1:p))
  {
    sum=(data[r,1]-yhat[r,])^2
    sum1=(yhat[r,]-mean(data[,1]))^2
    SSE=SSE+sum
    SSR=SSR+sum1
  }
  Rsqr[i]=(SSR/(SSE+SSR))*100
  MSE=SSE/p
  A=mx%*%C%*%t(mx)
  A1=(F-A)
  A2=(sum(diag(A1))/p)^2

```

```

    GCV[i]=MSE/A2
  }
GCV=as.matrix(GCV)
  Rsq=as.matrix(Rsq)
cat("=====
      =====", "\n")
cat("Nilai Knot dengan Spline linear 2 knot", "\n")

cat("=====
===== ", "\n")
print(knot2)
cat("=====
      =====", "\n")
cat("Rsq dengan Spline linear 2 knot", "\n")
cat("=====
      =====", "\n")
print (Rsq)
cat("=====
      =====", "\n")
cat("HASIL GCV dengan Spline linear 2 knot", "\n")
cat("=====
      =====", "\n")
print (GCV)
s1=min(GCV)
cat("=====
      =====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 2 knot", "\n")
cat("=====
      =====", "\n")
cat(" GCV =",s1, "\n")
write.table(GCV,file="d:/output GCV2.csv",sep=";")
write.table(Rsq,file="d:/output Rsq2.csv",sep=";")
write.table(knot2,file="d:/output knot2.csv",sep=";")
}

```

Lampiran 4. Program Pemilihan Titik Knot Optimal dengan Tiga Titik Knot Menggunakan *Software R*

```
GCV3=function(para)
{
  data=read.table("d://datatb.txt",header=FALSE)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-para-1
  F=matrix(0,nrow=p,ncol=p)
  dataA=data[, (para+2):q]
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot=matrix(ncol=m,nrow=nk)
  for (i in (1:m))
  {
    for (j in (1:nk))
    {
      a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
      knot[j,i]=a[j]
    }
  }
  knot=knot[2:(nk-1),]
  a2=nrow(knot)
  z=(a2*(a2-1)*(a2-2)/6)
  knot1=cbind(rep(NA,(z+1)))
  for (i in (1:m))
  {
    knot2=rbind(rep(NA,3))
    for (j in 1:(a2-2))
    {
      for (k in (j+1):(a2-1))
      {
        for (g in (k+1):a2)
        {
          xx=cbind(knot[j,i],knot[k,i],knot[g,i])
        }
      }
    }
  }
}
```

```

                                knot2=rbind(knot2,xx)
                                }
                        }
}
knot1=cbind(knot1,knot2)
}
knot1=knot1[2:(z+1),2:(3*m+1)]
aa=rep(1,p)
data1=matrix(ncol=(3*m),nrow=p)
data2=data[, (para+2):q]
a1=length(knot1[,1])
GCV=rep(NA,a1)
Rsqr=rep(NA,a1)
for (i in 1:a1)
{
  for (j in 1:ncol(knot1))
  {
    b=ceiling(j/3)
    for (k in 1:p)
    {
      if (data2[k,b]<knot1[i,j]) data1[k,j]=0 else
data1[k,j]=data2[k,b]-knot1[i,j]
    }
  }
}
mx=cbind(aa,data[,2:q],data1)
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in (1:p))
{
  sum=(data[r,1]-yhat[r,])^2
  sum1=(yhat[r,]-mean(data[,1]))^2
  SSE=SSE+sum
  SSR=SSR+sum1
}

```



```

}
Rsq[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}

GCV=as.matrix(GCV)
Rsq= as.matrix(Rsq)
cat("=====", "\n")
cat("Nilai Knot dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (knot1)
cat("=====", "\n")
cat("Rsq dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (Rsq)
r=max(Rsq)
print (r)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
cat(" GCV =",s1, "\n")
write.table(GCV,file="d:/output GCV3.csv",sep=";")
write.table(Rsq,file="d:/output Rsq3.csv",sep=";")
write.table(knot1,file="d:/output knot3.csv",sep=";")
}

```

Lampiran 5. Program Pemilihan Titik Knot Optimal dengan Kombinasi Titik Knot Menggunakan *Software R*

```
GCVkom=function(para)
{
  data=read.table("d://datatb.txt",header=FALSE)
  data=as.matrix(data)
  p1=length(data[,1])
  q1=length(data[1,])
  v=para+2
  F=matrix(0,nrow=p1,ncol=p1)
  diag(F)=1
  x1=read.table("d://x1.txt")
  x2=read.table("d://x2.txt")
  x3=read.table("d://x3.txt")
  x4=read.table("d://x4.txt")
  x5=read.table("d://x5.txt")
  n2=nrow(x1)
  a=matrix(nrow=5,ncol=3^5)
  m=0
  for (i in 1:3)
  for (j in 1:3)
  for (k in 1:3)
  for (l in 1:3)
  for (s in 1:3)
  {
    m=m+1
    a[,m]=c(i,j,k,l,s)
  }
  a=t(a)
  GCV=matrix(nrow=nrow(x1),ncol=3^5)
  for (i in 1:3^5)
  {
    for (h in 1:nrow(x1))
    {
      if (a[i,1]==1)
      {
```

```

gab=as.matrix(x1[,1])
gen=as.matrix(data[,v])
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,1]==2)
{
gab=as.matrix(x1[,2:3])
gen=as.matrix(cbind(data[,v],data[,v]))
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1: nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x1[,4:6])
gen=as.matrix(cbind(data[,v],data[,v],data[,v]))
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,2]==1)
{
gab=as.matrix(x2[,1] )
gen=as.matrix(data[,v+1])
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=1)

```

```

for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,2]==2)
{
gab=as.matrix(x2[,2:3] )
gen=as.matrix(cbind(data[, (v+1)],data[, (v+1)]))
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x2[,4:6])
gen=as.matrix(cbind(data[, (v+1)],data[, (v+1)],data[, (v+1)]))
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,3]==1)
{
gab=as.matrix(x3[,1] )
gen=as.matrix(data[, (v+2)])
cc=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{

```

```

if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,3]==2)
{
gab=as.matrix(x3[,2:3] )
gen=as.matrix(cbind(data[, (v+2)],data[, (v+2)]))
cc=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x3[,4:6])
gen=as.matrix(cbind(data[, (v+2)],data[, (v+2)],data[, (v+2)]))
cc=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,4]==1)
{
gab=as.matrix(x4[,1] )
gen=as.matrix(data[, (v+3)])
dd=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
}

```

```

else
if (a[i,4]==2)
{
gab=as.matrix(x4[,2:3] )
gen=as.matrix(cbind(data[, (v+3)], data[, (v+3)]))
dd=matrix(nrow=nrow(x1)*nrow(data), ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x4[,4:6])
gen=as.matrix(cbind(data[, (v+3)], data[, (v+3)], data[, (v+3)]))
dd=matrix(nrow=nrow(x1)*nrow(data), ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,5]==1)
{
gab=as.matrix(x5[,1] )
gen=as.matrix(data[, (v+4)])
ee=matrix(nrow=nrow(x1)*nrow(data), ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) ee[w,j]=0 else ee[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,5]==2)
{

```

```

gab=as.matrix(x5[,2:3] )
gen=as.matrix(cbind(data[, (v+4)], data[, (v+4)]))
ee=matrix(nrow=nrow(x1)*nrow(data), ncol=2)
for (j in 1:2)
  for (w in 1:nrow(data))
  {
    if (gen[w,j]<gab[h,j]) ee[w,j]=0 else ee[w,j]=gen[w,j]-gab[h,j]
  }
}
else
{
  gab=as.matrix(x5[,4:6])
  gen=as.matrix(cbind(data[, (v+4)], data[, (v+4)], data[, (v+4)]))
  ee=matrix(nrow=nrow(x1)*nrow(data), ncol=3)
  for (j in 1:3)
    for (w in 1:nrow(data))
    {

      if (gen[w,j]<gab[h,j]) ee[w,j]=0 else ee[w,j]=gen[w,j]-gab[h,j]
    }
  }
  ma=as.matrix(cbind(aa,bb,cc,dd,ee))
  mx=cbind(rep(1,nrow(data)), data[,2:q1], na.omit(ma))
  mx=as.matrix(mx)
  C=pinv(t(mx)%*%mx)
  B=C%*%(t(mx)%*%data[,1])
  yhat=mx%*%B
  SSE=0
  SSR=0
  for (r in 1:nrow(data))
  {
    sum=(data[r,1]-yhat[r,])^2
    sum1=(yhat[r,]-mean(data[,1]))^2
    SSE=SSE+sum
    SSR=SSR+sum1
  }
  Rsq=(SSR/(SSE+SSR))*100

```

```

MSE=SSE/p1
A=mx%%C%%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p1)^2
GCV[h,i]=MSE/A2
}

if (a[i,1]==1) sp=x1[,1] else
if (a[i,1]==2) sp=x1[,2:3] else
sp=x1[,4:6]
if (a[i,2]==1) spl=x2[,1] else
if (a[i,2]==2) spl=x2[,2:3] else
spl=x2[,4:6]
if (a[i,3]==1) splin=x3[,1] else
if (a[i,3]==2) splin=x3[,2:3] else
splin=x3[,4:6]
if (a[i,4]==1) spline=x4[,1] else
if (a[i,4]==2) spline=x4[,2:3] else
spline=x4[,4:6]
if (a[i,5]==1) splines=x5[,1] else
if (a[i,5]==2) splines=x5[,2:3] else
splines=x5[,4:6]
kkk=cbind(sp,spl,splin,spline,splines)
cat("=====", "\n")
print(i)
print(kkk)
print(Rsq)
}
write.table(GCV,file="d:/output GCV kombinasi.csv",sep=";")
write.table(Rsq,file="d:/output Rsq kombinasi.csv",sep=";")
}

```


Lampiran 6. Program Kombinasi Titik Knot 3^5

```
GCVkom=function(data,para)
{
a=matrix(nrow=5,ncol=3^5)
m=0
for (i in 1:3)
for (j in 1:3)
for (k in 1:3)
for (l in 1:3)
for (n in 1:3)

{
m=m+1
a[,m]=c(i,j,k,l,n)
print(a[,m])
}
}
```

Lampiran 7. Program Estimasi Parameter dengan Kombinasi Titik Knot 1,3,2,3,2

```

uji=function(alpha,para)
{
data=read.table("d:/datatb.txt")
knot=read.table("d:/knot.txt")
data=as.matrix(data)
knot=as.matrix(knot)
ybar=mean(data[,1])
m=para+2
p=nrow(data)
q=ncol(data)
dataA=cbind(data[,m],data[,m+1],data[,m+1],data[,m+1],data[,m
+2],data[,m+2],data[,m+3],data[,m+3],data[,m+3],data[,m+4],dat
a[,m+4])
dataA=as.matrix(dataA)
satu=rep(1,p)
n1=ncol(knot)
data.knot=matrix(ncol=n1,nrow=p)
for (i in 1:n1)
{
for(j in 1:p)
{
if (dataA[j,i]<knot[1,i]) data.knot[j,i]=0 else
data.knot[j,i]=dataA[j,i]-knot[1,i]
}
}
mx=cbind(satu,data[,2],data.knot[,1],data[,3],data.knot[,2:4],data[
,4],data.knot[,5:6],data[,5],data.knot[,7:9],data[,6],data.knot[,10:1
1])
mx=as.matrix(mx)
B=(pinv(t(mx)%*%mx))%*%t(mx)%*%data[,1]
cat("=====", "\n")
cat("Estimasi Parameter", "\n")
cat("=====", "\n")
print (B)

```

```

n1=nrow(B)
yhat=mx%*%B
res=data[,1]-yhat
SSE=sum((data[,1]-yhat)^2)
SSR=sum((yhat-ybar)^2)
SST=SSR+SSE

MSE=SSE/(p-n1)
MSR=SSR/(n1-1)
Rsqr=(SSR/(SSR+SSE))*100

#uji F (uji serentak)
Fhit=MSR/MSE
pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
if (pvalue<=alpha)
{
  cat("-----","\n")
  cat("Kesimpulan hasil uji serentak","\n")
  cat("-----","\n")
  cat("Tolak Ho yakni minimal terdapat 1 prediktor yang
signifikan","\n")
  cat("", "\n")
}
else
{
  cat("-----","\n")
  cat("Kesimpulan hasil uji serentak","\n")
  cat("-----","\n")
  cat("Gagal Tolak Ho yakni semua prediktor tidak berpengaruh
signifikan","\n")
  cat("", "\n")
}

#uji t (uji individu)
thit=rep(NA,n1)
pval=rep(NA,n1)
SE=sqrt(diag(MSE*(pinv(t(mx)%*%mx))))

```

```

cat("-----","\n")
cat("Kesimpulan hasil uji individu","\n")
cat("-----","\n")
thit=rep(NA,n1)
pval=rep(NA,n1)
for (i in 1:n1)
{
thit[i]=B[i,1]/SE[i]
pval[i]=2*(pt(abs(thit[i]),(p-n1),lower.tail=FALSE))
if (pval[i]<=alpha) cat("Tolak Ho yakni prediktor signifikan
dengan pvalue",pval[i],"\\n") else cat("Gagal tolak Ho yakni
prediktor tidak signifikan dengan pvalue",pval[i],"\\n")
}
}
thit=as.matrix(thit)
cat("=====","\n")
cat("nilai t hitung","\n")
cat("=====","\n")
print (thit)
cat("Analysis of Variance","\n")
cat("=====","\n")
cat("Sumber      df    SS    MS    Fhit","\n")
cat("Regresi      ",(n1-1)," ",SSR," ",MSR,"",Fhit,"\\n")
cat("Error        ",p-n1," ",SSE,"",MSE,"\\n")
cat("Total        ",p-1," ",SST,"\\n")
cat("=====","\n")
cat("s=",sqrt(MSE),"      Rsq=",Rsq,"\\n")
cat("pvalue(F)=",pvalue,"\\n")
write.table(res,file="d:/output uji residual knot3.csv",sep=";")
write.table(pval,file="d:/output uji pvalue knot3.csv",sep=";")
write.table(mx,file="d:/output uji mx knot3.csv",sep=";")
write.table(yhat,file="d:/output uji yhat knot3.csv",sep=";")
}

```

Lampiran 8. Program Uji Glejser untuk Kombinasi Titik Knot 1,3,2,3,2

```
glejser=function(data,knot,res,alpha,para)
{
data=read.table("d:/datatb.txt")
knot=read.table("d:/knot.txt")
res=read.table("e:/resi.txt")
data=as.matrix(data)
knot=as.matrix(knot)
res=abs(res)
res=as.matrix(res)
rbar=mean(res)
m=para+2
p=nrow(data)
q=ncol(data)
dataA=
cbind(data[,m],data[,m+1],data[,m+1],data[,m+1],data[,m+2],data
[,m+2],data[,m+3],data[,m+3],data[,m+3],data[,m+4],data[,m+4])
dataA=as.matrix(dataA)
satu=rep(1,p)
n1=ncol(knot)
data.knot=matrix(ncol=n1,nrow=p)
for (i in 1:n1)
{
for(j in 1:p)
{
if (dataA[j,i]<knot[1,i]) data.knot[j,i]=0 else
data.knot[j,i]=dataA[j,i]-knot[1,i]
}
}
mx=
cbind(satu,data[,2],data.knot[,1],data[,3],data.knot[,2:4],data[,4],d
ata.knot[,5:6],data[,5],data.knot[,7:9],data[,6],data.knot[,10:11])
mx=as.matrix(mx)
B=(ginv(t(mx)%*%mx))%*%t(mx)%*%res
n1=nrow(B)
yhat=mx%*%B
```

```

residual=res-yhat
SSE=sum((res-yhat)^2)
SSR=sum((yhat-rbar)^2)
SST=SSR+SSE
MSE=SSE/(p-n1)
MSR=SSR/(n1-1)
Rsqr=(SSR/SST)*100
#uji F (uji serentak)
Fhit=MSR/MSE
pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
if (pvalue<=alpha)
{
cat("-----","\n")
cat("Kesimpulan hasil uji serentak","\n")
cat("-----","\n")
cat("Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan
atau terjadi heteroskedastisitas","\n")
cat("","\n")
}
else
{
cat("-----","\n")
cat("Kesimpulan hasil uji serentak","\n")
cat("-----","\n")
cat("Gagal Tolak Ho yakni semua prediktor tidak berpengaruh
signifikan atau tidak terjadi heteroskedastisitas","\n")
cat("","\n")
}
cat("Analysis of Variance","\n")
cat("=====","","\n")
cat("Sumber    df    SS    MS    Fhit","\n")
cat("Regresi    ",(n1-1)," ",SSR," ",MSR,"",Fhit,"\n")
cat("Error      ",p-n1," ",SSE,"",MSE,"\n")
cat("Total      ",p-1," ",SST,"\n")
cat("=====","","\n")
cat("s=",sqrt(MSE),"    Rsqr=",Rsqr,"\n")
cat("pvalue(F)=",pvalue,"\n")
}

```

Lampiran 9. Output Nilai GCV dengan Satu Titik Knot

No	GCV	$X1$	$X2$	$X3$	$X4$	$X5$
1	840374,4	34,96224	5,022857	0,083673	2,99902	21,05204
2	840229,2	36,22449	5,455714	0,167347	3,277041	22,04408
3	846474,8	37,48673	5,888571	0,25102	3,555061	23,03612
4	804860,9	38,74898	6,321429	0,334694	3,833082	24,02816
5	766596	40,01122	6,754286	0,418367	4,111102	25,0202
6	746074	41,27347	7,187143	0,502041	4,389122	26,01224
7	743255,9	42,53571	7,62	0,585714	4,667143	27,00429
8	749616,3	43,79796	8,052857	0,669388	4,945163	27,99633
9	755745,9	45,0602	8,485714	0,753061	5,223184	28,98837
10	759499,3	46,32245	8,918571	0,836735	5,501204	29,98041
⋮	⋮	⋮	⋮	⋮	⋮	⋮
48	692986,7	94,28776	25,36714	4,016327	16,06598	67,67796

Lampiran 10. Output Nilai GCV dengan Dua Titik Knot

No	GCV	$X1$	$X2$	$X3$	$X4$	$X5$
1	840374,3629	33,7	4,59	0	2,721	20,06
		34,9622	5,0228	0,0836	2,9990	21,0520
2	840229,1859	33,7	4,59	0	2,721	20,06
		36,2244	5,4557	0,1673	3,2770	22,0440
3	846474,7947	33,7	4,59	0	2,721	20,06
		37,4867	5,888	0,2510	3,5550	23,0361
4	804860,8891	33,7	4,59	0	2,721	20,06
		38,7489	6,3214	0,3346	3,8330	24,0281
5	766595,9560	33,7	4,59	0	2,721	20,06
		40,0112	6,7542	0,4183	4,1111	25,020
6	746074,0011	33,7	4,59	0	2,721	20,06
		41,2734	7,1871	0,5020	4,3891	26,0122
7	743255,9430	33,7	4,59	0	2,721	20,06
		42,5357	7,62	0,5857	4,6671	27,0042
8	749616,2592	33,7	4,59	0	2,721	20,06
		43,7979	8,0528	0,6693	4,9451	27,996
⋮	⋮	⋮	⋮	⋮	⋮	⋮
1225	692986,7198	94,287	25,367	4,0163	16,065	67,677
		95,55	25,8	4,1	16,344	68,67

Lampiran 11. Output Nilai GCV dengan Tiga Titik Knot

No	GCV	x1	x2	x3	x4	x5
1	644599,4138	34,96224	5,022857	0,08367	2,9990	21,05204
		36,22449	5,455714	0,16734	3,2770	22,04408
		37,48673	5,888571	0,25102	3,5550	23,03612
2	606792,3071	34,96224	5,022857	0,08367	2,9990	21,05204
		36,22449	5,455714	0,16734	3,2770	22,04408
		38,74898	6,321429	0,33469	3,8330	24,02816
3	709486,3087	34,96224	5,022857	0,08367	2,9990	21,05204
		36,22449	5,455714	0,16734	3,2770	22,04408
		40,01122	6,754286	0,41836	4,1111	25,0202
4	799607,5745	34,96224	5,022857	0,08367	2,9990	21,05204
		36,22449	5,455714	0,16734	3,2770	22,04408
		41,27347	7,187143	0,50204	4,3891	26,01224
⋮	⋮	⋮	⋮	⋮	⋮	⋮
17296	530760,3761	91,76327	24,50143	3,84898	15,509	65,69388
		93,02551	24,93429	3,93265	15,787	66,68592
		94,28776	25,36714	4,01632	16,065	67,67796

Lampiran 12. Output Estimasi Parameter dan Uji Signifikansi Parameter Model

Estimasi Parameter

```

[1,] 2184.864056
[2,] -2.564361
[3,] -75.863216
[4,] 141.526198
[5,] -128.430085
[6,] -21.847054
[7,] -10.923653
[8,] 176.203967
[9,] -6129.817085
[10,] -2043.264222
[11,] -595.473536
[12,] 649.170892
[13,] -368.316391
[14,] -184.158195
[15,] 12.707361
[16,] 2647.442589
[17,] -7652.740754

```

Kesimpulan hasil uji serentak

Tolak H_0 yakni minimal terdapat 1 prediktor yang signifikan

Kesimpulan hasil uji individu

Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue
0.05915988

Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue
0.7793626

Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue
0.2688836

Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.05166073
 Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.1209485
 Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.9664244
 Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.9664244
 Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.1426633
 Tolak H_0 yakni prediktor signifikan dengan pvalue 0.005388391
 Tolak H_0 yakni prediktor signifikan dengan pvalue 0.005388391

 Tolak H_0 yakni prediktor signifikan dengan pvalue $9.020632e-05$
 Tolak H_0 yakni prediktor signifikan dengan pvalue 0.0003451616
 Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.6015729
 Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.6015729
 Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.06177257
 Tolak H_0 yakni prediktor signifikan dengan pvalue $1.0009e-05$
 Tolak H_0 yakni prediktor signifikan dengan pvalue $1.498547e-05$

nilai t hitung

[,1]
 [1,] 1.99518718
 [2,] -0.28377388
 [3,] -1.13568625
 [4,] 2.06333110
 [5,] -1.61632538
 [6,] -0.04259805
 [7,] -0.04259805
 [8,] 1.52305267
 [9,] -3.10278044

[10,] -3.10278044
 [11,] -4.82729856
 [12,] 4.26453351
 [13,] -0.53014464
 [14,] -0.53014464
 [15,] 1.97324937
 [16,] 5.76871612
 [17,] -5.59322919

Analysis of Variance

Sumber	df	SS	MS	Fhit
Regresi	16	21078174	1317386	6.602783
Error	21	4189916	199519.8	
Total	37	25268090		

s= 446.6764 Rsq= 83.41815

pvalue(F)= 5.094488e-05

Lampiran 13. Output Uji Glejser

Kesimpulan hasil uji serentak

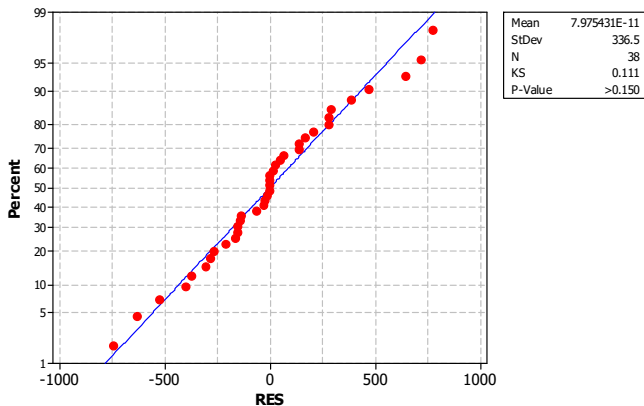
Gagal Tolak H_0 yakni semua prediktor tidak berpengaruh signifikan atau tidak terjadi heteroskedastisitas

Analysis of Variance

Sumber	df	SS	MS	Fhit
Regresi	16	827030.4	51689.4	0.9830687
Error	21	1104173	52579.64	
Total	37	1931203		

$s = 229.3025$ $Rsq = 42.82462$
 $pvalue(F) = 0.5058166$

Lampiran 14. Output Uji *Kolmogorov-Smirnov*



Lampiran 15. Surat Pernyataan Data Sekunder**SURAT KETERANGAN**

Saya yang bertanda tangan di bawah ini menerangkan bahwa :

- 1 Mahasiswa Statistika FMIPA – ITS dengan identitas berikut :

Nama : FIRDA FAHRUN NISA'

NRP : 1312100027

Telah mengambil data di Instansi / perusahaan kami :

Nama Instansi : Dinas Kesehatan Provinsi Jawa Timur

Divisi / Bagian : Seksi Informasi Litbangkes

Sejak tanggal 19 Januari 2016 sampai dengan 30 Juni 2016 untuk keperluan Tugas Akhir Semester Genap 2015/2016.


- 2 Tidak keberatan / ~~Keberatan~~* nama Instansi di cantumkan dalam Tugas Akhir Mahasiswa Statistika yang akan di simpan di Perpustakaan ITS dan dibaca di lingkungan ITS .
- 3 Tidak keberatan / ~~Keberatan~~* bahwa hasil analisis data dari Instansi Perusahaan dipublikasikan dalam E Journal ITS yaitu Jurnal Sains dan Seni ITS.

SURABAYA, Juni 2016

KEPALA SEKSI INFORMASI DAN LITBANGKES

BIDANG PPKM

DINAS KESEHATAN PROVINSI JAWA TIMUR



DESI AVIAJIATI, SKM ,M.K es

NIP.19601206 198303 2 011

*(coret yang tidak perlu)

BIODATA PENULIS



Penulis dengan nama lengkap Firda Fahrudin Nisa' dan nama panggilan Firda dengan tempat tanggal lahir Lumajang, 7 Mei 1994. Bungsu dari dua bersaudara pasangan Bapak Sukastono dan Ibu Rochimi. Penulis memiliki hobi mendengarkan musik, menulis, menyanyi, dan menonton film. Pendidikan formal yang telah ditempuh selama 12 tahun oleh penulis SD, SMP, dan SMA Taruna Dra. Zulaeha

Leces, Probolinggo. Pada tahun 2012 penulis diterima menjadi mahasiswa Jurusan Statistika ITS melalui jalur undangan dengan NRP 1312100027. Semasa kuliah yang ditempuh dalam 4 tahun, penulis aktif di organisasi kemahasiswaan ITS di tingkat jurusan yaitu Himpunan Mahasiswa Statistika (HIMASTA-ITS) pada periode 2013-2014 sebagai *staff* di Departemen Pengembangan Sumber Daya Mahasiswa (PSDM) dan pada periode 2014-2015 sebagai *Steering Committee* (SC) Bina Cinta Statistika (BCS) 2014 dan *Steering Committee* (SC) GERIGI ITS 2014. Selain itu penulis pernah turut berpartisipasi dalam komunitas TEDxITS sebagai *Liaison Officer*. Segala kritik dan saran serta diskusi lebih lanjut mengenai Tugas Akhir ini dapat melalui *e-mail* di firdafahrudin@gmail.com.